

EXHIBIT 1

JOINT LETTER BRIEF Docket 247

October 31, 2024

The Honorable Thomas S. Hixson
United States District Court for the Northern District of California
San Francisco Courthouse, Courtroom E – 15th Floor
450 Golden Gate Avenue
San Francisco, CA 94102

Re: Kadrey et al v. Meta Platforms, Inc., No. 3:23-cv-03417-VC

Dear Magistrate Judge Hixson:

Plaintiffs and Meta Platforms, Inc. (“Meta”) respectfully submit this joint letter brief in response to the Court’s October 29, 2024 Discovery Order (Dkt. 243; the “Discovery Order”). The Parties are pleased to report that they have resolved a number of open issues identified by the Court in the Discovery Order. This letter brief begins by describing the status on those resolved issues, and then sets out the Parties’ respective positions on the issues remaining for the Court.

I. Deposition Dates: The parties have agreed to dates for the depositions of Alex Boesenberg (November 18) and Amrish Acharya (November 21). With respect to Mark Zuckerberg’s deposition, immediately after receipt of the Court’s Order, Meta’s counsel sent Plaintiffs’ counsel a draft stipulation for submission to Judge Chhabria, which – as provided by the Parties’ agreement – would allow this deposition to proceed on December 17, four days after the December 13, 2024 cutoff. That stipulation is attached as **Exhibit A** to the Declaration of Judd Lauter (“Lauter Decl.”). The parties hope to finalize and file the stipulation immediately and no later than this week.

II. Meta’s 30(b)(6) Designees: The Parties met and conferred regarding Plaintiffs’ thirteen amended 30(b)(6) topics and Meta’s objections thereto on October 29. Although certain issues remain for resolution by the Court as to five topics (discussed below), the Parties were able to resolve all disputes with respect to eight topics. Moreover, Meta has provided its designees for each of these topics to Plaintiffs, and the depositions are scheduled as follows: Topics 1–3, 6 (Mike Clark; November 14), Topics 7, 10 (Nikolay Bashlykov; December 6), Topic 9 (Amrish Acharya; November 21), Topic 13 (Eugene Nho; December 6).

III. Disputed Issues: The Parties respectfully seek the Court’s guidance on four open issues.

A. Plaintiffs’ Request to Depose Chris Cox / Meta’s Motion for Protective Order

Meta’s Position:¹

Soon after Judge Chhabria “reluctantly” extended the case schedule by roughly two months, ECF No. 211, this Court sought the parties’ “proposals for what the deposition limits should be,” including “specific witnesses and topics each side believes are appropriate for deposition,” ECF No. 213 (emphasis added). Plaintiffs submitted their proposal on October 11, 2024, *see* ECF No. 229, naming dozens of Meta employees and third parties whose depositions they claimed to need as part of their request for 31 additional depositions, *see id.* at 4–6. Notably absent from Plaintiffs’ list was Mr. Cox, Meta’s Chief Product Officer, whom Plaintiffs now seek to depose despite that he is an “apex” witness with no unique knowledge or information to provide. The Court should bar Mr. Cox’s deposition, or, alternatively, limit the deposition to two hours.

Plaintiffs listed Mr. Cox in their December 2023 initial disclosures, meaning they have known his identity for at least 10 months. However, Plaintiffs expressed no interest in taking his deposition until two weeks ago. In prior submissions to the Court, Plaintiffs identified numerous witnesses whose depositions were – according to Plaintiffs – important to their case. For example, in their initial September 12, 2024 request for additional depositions prior to Judge Chhabria’s extension of the fact discovery cutoff, Plaintiffs listed 28 proposed witnesses they claimed they needed to depose to develop their case for summary judgment; *Mr. Cox was not among them*. ECF No. 143. Then, on October 15, 2024, after Judge Chhabria extended the fact discovery cutoff, Plaintiffs sent an even longer list of proposed deponents; again, Mr. Cox was not included. ECF No. 229. It was only after this Court partially granted Plaintiffs’ request for additional depositions on October 15, 2024 (ECF No. 231) that Plaintiffs for the first time sought a date for Mr. Cox’s deposition.

At Meta’s request, the parties met and conferred on October 29, 2024 regarding Plaintiffs’ basis for seeking Mr. Cox’s deposition. During that call, Plaintiffs did not dispute that Mr. Cox is an “apex” witness. Rather, Plaintiffs argued that Mr. Cox’s deposition is warranted because he supposedly has first-hand knowledge of relevant facts and disputed that a showing of unique knowledge was required. In support of their position, Plaintiffs cited a *New York Times* article that Plaintiffs previously presented to the Court in September (ECF No. 143). That article alleges – based on an unidentified source – that Mr. Cox participated in a purported discussion with other

¹ Meta has endeavored to address all of Plaintiffs’ arguments in each of the sections of the letter brief, in which Plaintiffs have chosen to focus more on *document* discovery than deposition issues. Any lack of response by Meta to a particular argument should not be construed as conceding or agreeing to Plaintiffs’ assertions.

Meta senior executives concerning the use of copyrighted data to train its LLMs. During the meet and confer, Plaintiffs also cited Mr. Cox’s alleged knowledge of and involvement in the acquisition of data for training LLMs. However, these examples show—at most—that Mr. Cox was one of many participants in discussions about topics related to this case; they do not reflect that Mr. Cox has unique knowledge or that the information sought by Plaintiffs cannot be obtained by less intrusive methods.²

“When a party seeks the deposition of a high-level executive (a so-called “apex” deposition), courts have observed that such discovery creates a tremendous potential for abuse or harassment.” *Apple Inc. v. Samsung Elecs. Co., Ltd.*, 282 F.R.D. 259, 263 (N.D. Cal. 2012) (cleaned up); *see, e.g., Davis v. Pinterest, Inc.*, 2021 WL 11117688, at *1 (N.D. Cal. May 27, 2021). “The court therefore has discretion to limit discovery where the discovery sought ‘can be obtained from some other source that is more convenient, less burdensome, or less expensive’” *Id.* “In determining whether to allow an apex deposition, courts consider (1) whether the deponent has unique first-hand, non-repetitive knowledge of the facts at issue in the case and (2) whether the party seeking the deposition has exhausted other less intrusive discovery methods.” *Samsung*, 282 F.R.D. at 263 (cleaned up); *see also Finisar Corp. v. Nistica, Inc.*, 2015 WL 3988132, *1 (N.D. Cal. Jun. 30, 2015) (setting forth this framework). Here, these two prongs weigh heavily against allowing Mr. Cox’s deposition, because he has no unique, firsthand, non-cumulative knowledge, and because Plaintiffs have many other witnesses – including Meta’s CEO and other members of Meta’s leadership team – with whom they can cover the same ground as well as other forms of discovery to obtain the information sought.

Plaintiffs, for the first time in this submission, dispute that Mr. Cox qualifies as an apex witness. Mr. Cox is Meta’s Chief Product Officer, a C-level executive of the company listed among Meta’s five senior-most “Executives” – along with Meta’s CEO, President of Global Affairs, Chief Financial Officer, and Chief Operating Officer.³ Contrary to Plaintiffs’ assertion, courts have considered executives with similar titles as apex witnesses. *See, e.g., Symantec Corp. v. Acronis, Inc.*, 2013 WL 12184286, at *1 (N.D. Cal. Oct. 10, 2013) (considering “President of Products and Services” an apex witness). Like the witness in *Symantec*, there is no legitimate dispute that Mr. Cox is an apex witness.

As stated above, Mr. Cox is Meta’s Chief Product Officer, but Plaintiffs are already set to depose two of Meta’s Directors of Product Management—Mike Clark and Chaya Nayak. Both

² Plaintiffs mischaracterize this article as suggesting that Mr. Cox may have unique information about a one-on-one conversation he supposedly had with an “unnamed Meta employee” about the use of copyrighted data to train Meta’s models. The article says no such thing; it provides a third-hand hearsay account of a *speculative* discussion among Mr. Cox and other senior executives. Plaintiffs have not even come forward with evidence that this conversation occurred, much less that they cannot obtain all relevant evidence about Meta’s decisionmaking around sources of training data from other witnesses they will depose.

³ <https://about.meta.com/media-gallery/executives/>.

ultimately report to Mr. Cox, are closer to the issues, and can testify about product-related topics, obviating any need for Mr. Cox's duplicative testimony. *See Symantec*, 2013 WL 12184286, at *1 (denying motion to compel deposition of Plaintiff's President of Products and Services in part because Defendants previously deposed a Vice President of Product Management on related topics). To the extent Plaintiffs seek the perspectives of Meta's leadership, they already deposed Ahmad Al-Dahle (VP of Meta's GenAI Division) and will be deposing Mark Zuckerberg (Meta's Founder and CEO). Notably, Mr. Cox appears in fewer than 100 of the over 20,000 Meta documents produced to date.

Below, Plaintiffs suggest that this number is low because Mr. Cox was not a custodian, but if he actually had the "pivotal role" Plaintiffs seek to ascribe to him, one would expect to see him on more emails and documents sent and received by designated custodians, such as Mr. Al-Dahle and Directors of Product Management Mike Clark and Chaya Nayak. Further, Plaintiffs' vague assertion that Meta did not adequately search for and produce documents for its custodians from all appropriate data sources is both unsupported and untimely. *See, infra* Section C (Meta's Position). Plaintiffs even go so far as to fault Meta for its supposed failure to "create ... relevant documents" about the facts and events they speculate might have occurred. None of this justifies the burdensome fishing expedition Plaintiffs seek.

Plaintiffs further argue that these documents show Mr. Cox's involvement in discussions and decision-making around data acquisition for LLMs and place significant emphasis on the LibGen database, which is not mentioned in the operative Complaint and does not form the basis for Plaintiffs' claims here. Based on counsel's review, other than isolated instances, he is mentioned only in passing or alongside other Meta employees and executives that Plaintiffs have deposed or are planning to depose. The documents Plaintiffs cite support this: Meta_Kadrey_00048202⁴ refers to Mr. Cox, Mr. Zuckerberg, whom Plaintiffs will depose, and Andrew Bosworth, Meta's Chief Technology Officer, as a unit ("Boz/Zuck/Cox"); Meta_Kadrey_00088838-9 similarly refers to Mr. Cox alongside Mr. Zuckerberg and Mr. Al-Dahle ("zuck/cox/ahmad") and concludes that Mr. Cox was *not* needed to review the decisions discussed; and Meta_Kadrey_0044964 lists Mr. Cox alongside Nick Clegg, Meta's President of Global Affairs merely as "spokespeople." *See Affinity Labs of Texas v. Apple, Inc.*, No., 2011 WL 1753982, 2011 WL 1753982, at *16 ("Courts have repeatedly denied apex depositions even on a showing that the executive made public statements on relevant issues."). The remaining documents cited by Plaintiffs show Mr. Cox's high-level involvement in "data strategy," generally, and among other Meta witnesses that will be or have been deposed. And, as noted, the *New York Times* article that Plaintiffs cite as justifying Mr. Cox's deposition mentions him only once in passing and merely as one alleged participant in a discussion among senior executives. This record does not suggest that Mr. Cox has unique, firsthand knowledge that cannot be obtained through

⁴ This document also concerns the use of video data, which is not at issue in this dispute, rather than text data.

the depositions of other Meta witnesses already scheduled (which is presumably why Plaintiffs repeatedly requested those other witnesses' depositions and not that of Mr. Cox).

Notably, Plaintiffs may obtain and have obtained the information they allegedly seek from Mr. Cox from other sources. This includes ESI and document discovery served in the case, including new written discovery served in mid-October. It also includes prior and scheduled depositions of numerous other Meta witnesses, including various senior-level personnel such as Mark Zuckerberg (Meta's Founder and CEO), Ahmad Al-Dahle, (VP of Meta's GenAI Division), Joelle Pineau (Meta's VP of AI Research), Sergey Edunov (Meta's Director of AI Research), and Yann LeCun (Meta VP and Chief AI Scientist) – in addition to Mr. Clark and Ms. Nayak. Thus, the deposition of Mr. Cox is “unreasonably cumulative, or can be obtained from some other source that is more convenient, less burdensome, or less expensive.” Fed. R. Civ. P. 26(b)(2)(C)(i). Indeed, Plaintiffs' own approach demonstrates that taking Mr. Cox's deposition was an afterthought, given that he did not appear on any of the prior lists of witnesses Plaintiffs felt necessary to depose. *Compare* ECF No. 143 and 226.

Plaintiffs claim that the Court should allow the deposition of Mr. Cox because other high level executives (such as Mr. Zuckerberg and Mr. Al-Dahle) have been or will be deposed and Mr. Cox should not receive “more favorable treatment,” but this badly misses the point. Having been granted the right to depose Mr. Zuckerberg and others on the facts they claim are central to this case, Plaintiffs cannot show any legitimate need to also depose Mr. Cox on those same matters.

Although Meta submits that the Court should issue a protective order preventing Mr. Cox's deposition, in the alternative, if the Court is inclined to allow Mr. Cox's deposition, Meta respectfully requests that the Court limit Mr. Cox's deposition to two hours. Under Federal Rule 26(c)(1), courts frequently limit the length of an “apex” deposition to no more than three hours. *See, e.g., Apple Inc. v. Samsung Elecs. Co., Ltd.*, 282 F.R.D. 259, 265–69 (N.D. Cal. 2012) (limiting depositions to two or three hours each); *Finisar Corp. v. Nistica, Inc.*, 2015 WL 3988132, at *4 (N.D. Cal. June 30, 2015) (two hours); *Hunt v. Cont'l Casualty Co.*, 2015 WL 1518067, *3 (N.D. Cal. Apr. 3, 2015) (three hours); *In re Transpacific Passenger Air Transp. Antitrust Litig.*, 2014 WL 939287, at *1 (N.D. Cal. Mar. 6, 2014) (two hours); *In re Google Litig.*, 2011 WL 4985279, at *2 (N.D. Cal. Oct. 19, 2011) (three hours); *Google Inc. v. Am. Blind & Wallpaper Factory, Inc.*, 2006 WL 2578277, at *3 (N.D. Cal. Sept. 6, 2006), objections overruled, 2006 WL 3050866 (N.D. Cal. Oct. 23, 2006) (three hours). The appropriate length of an apex deposition depends on the extent of the deponent's unique, firsthand, relevant knowledge, as well as the exhaustion of other discovery methods. *See, e.g., Samsung*, 282 F.R.D. at 263–64.

Here, two hours on the record will provide Plaintiffs with ample time to cover the limited issues they have identified as justifying his deposition.⁵ Limiting Mr. Cox's deposition to two

⁵ Unlike the named Plaintiffs in this case, who are the primary sources of information relating to their claims, Plaintiffs can obtain the information they seek from Mr. Cox from other deponents.

hours is particularly appropriate given the high-level nature of his involvement in Meta's AI initiatives (as reflected in Plaintiffs' cited documents) and Plaintiffs' ability to obtain information they seek from their discovery requests and Meta witnesses that Plaintiffs have deposed or will depose such as the senior-level personnel and other witnesses referenced above, including Messrs. Zuckerberg, Al-Dahle, LeCun, and Edunov and Ms. Pineau. Given others' knowledge about the limited topics about which Plaintiffs seek to question Mr. Cox, Plaintiffs do not need seven hours to question Mr. Cox about those same topics.

Plaintiffs' Position:

A. Plaintiff's Position Regarding Deposition Testimony from Chris Cox

This Court has already denied an apex gambit by Meta, holding that, in this important case, the apex doctrine does not shield Meta's CEO, Mark Zuckerberg, from testifying. Chris Cox, who holds an inferior position to Mr. Zuckerberg, is not entitled to more favorable treatment. Mr. Cox is Meta's "Chief Product Officer." Meta has not shown that makes him a so-called "apex witness," but even if it does, one thing is clear: as the head of "product," he is positioned to address the central issue of why Meta chose to use pirated data sources to train its AI. Training on "shadow libraries" like LibGen, a database developed over decades by Russian hackers, posed great risks to Meta. Cox may be the person, or one of a select group, who decided to run that risk rather than negotiate licenses to acquire data in compliance with copyright law. Indeed, the documents and testimony already produced suggest that Mr. Cox has firsthand knowledge of the issues at the heart of this case. And, as this Court has already decided, that is all that is needed to take his deposition, absent an extraordinary showing by Meta.

1. Meta Has Not Made the Extraordinary Showing Necessary to Bar Mr. Cox's Testimony

In this very case, the Court has explained that apex depositions "'should be allowed'" "where a corporate officer may have any first-hand knowledge of relevant facts." Dkt. 171 at 1–2 (quoting *Finisar Corporation v. Nistica, Inc.*, 2015 WL 3988132, *2 (N.D. Cal. June 30, 2015)). Further, "formal exhaustion" of every other potential source of discovery is "not 'viewed as an absolute requirement.'" *Id.* at 2 (quoting *Hunt v. Continental Casualty Co.*, 2015 WL 1518067, *2 (N.D. Cal. April 3, 2015)). Instead, a court will bar the deposition only upon a showing of "'extraordinary circumstances.'" *Id.* (quoting *Powertech Technology, Inc., Tessera, Inc.*, 2013 WL 3884254 (N.D. Cal. July 26, 2013)). Meta has not come close to overcoming its "extraordinary" burden.

A. As this Court has held, a party must "provide[] evidence to substantiate its apex argument; job titles are not enough." *Trustlabs, Inc. v. Jaiyong*, No. 21-02606, 2023 WL 8230559, at *4 (N.D. Cal. June 26, 2023). Meta makes no showing that Mr. Cox qualifies as an "apex" witness, stating only his job title and that he is a "high-level executive." But Meta has dozens of "high-level executives," listing 24 of them on its website, including executives whom it does not claim are "apex" (like Mr. Al-Dahle).⁶ Plaintiffs are aware of no case, and Meta cites none, in which a

⁶ <https://about.meta.com/media-gallery/executives/>

“chief product officer” has been shielded from testifying under the apex doctrine. For this reason alone, Meta’s motion should be denied.

B. Meta does not contest that Mr. Cox has highly relevant, first-hand knowledge. And he surely does, based on publicly available information and documents and testimony already produced in this case. Mr. Cox appears to have been the key driver of Meta’s AI data strategy and the main decision-maker on many data-specific decisions. To begin with, Mr. Cox was listed as a Meta representative who can deny or approve requests for budget increases for what appears to be purchase of training data. *See* Ex. A, Meta_Kadrey_00048202. He was also documented requesting from his subordinates a “data strategy” for “Beyond Llama 4,” suggesting he was pivotal to Meta’s judgment calls pertaining to AI training data. Ex B, Meta_Kadrey_00080096. On that same note, Mr. Cox was mentioned in documents indicating that he was needed to review decisions related to acquisition of data. Ex. C, Meta_Kadrey_00088838-9. He was also listed as a key decision-maker at recurring meetings to discuss development of LLMs. Ex. D, Meta_Kadrey_00093952. Moreover, Mr. Cox’s role between Messrs. Al-Dahle and Zuckerberg on the organizational chart, Ex. E, Al-Dahle Dep. Tr. 139:13-14, suggests that he had the final or near-final say on many critical calls that are at issue in this case. *See also* Ex. F, Meta_Kadrey_0044964 (listing him one of two “Business & Policy Elites” along with Nick Clegg).

As an example of his pivotal role, Mr. Cox led a key meeting that may define Meta’s state of mind in its use of copyrighted material, as well as the commercial purpose behind the same. As detailed in a *New York Times* article, ***How Tech Giants Cut Corners to Harvest Data for A.I.***, that meeting involved Mr. Cox and an unnamed Meta employee, and no records of it have been produced. According to that article, Mr. Cox had a “separate discussion” with an employee about the use of copyrighted data to train Meta’s models, suggesting he has highly relevant—and perhaps unique—information. Ex. G, Kadrey_00037448. Meta notably does not claim the article’s account is incorrect. Nor does it claim that another witness can testify as to what occurred at this meeting.⁷

In short, there is “evidence of [Mr. Cox’s] specific involvement in the company’s AI initiatives.” Dkt. 171 at 2. There is evidence that Mr. Cox was a “principal decision maker” on AI data sets. *Id.* And there is “evidence of [Mr. Cox’s] direct supervision of Meta’s AI products.” *Id.* “This showing justifies an apex deposition.” *Id.*

C. Rather than contest Mr. Cox’s relevance, Meta’s entire argument is that deposing Mr. Cox would be “duplicative” and the information he possesses may be “obtained ... from other sources.” But, as Meta concedes, Plaintiffs held off on seeking Mr. Cox’s deposition until many other sources of discovery had been pursued. But that discovery has left open critical questions, such as why Meta chose to use hacked book datasets rather than pay for licenses. Perhaps intentionally, Meta seems to have created little documentation of its decisions (or is improperly withholding them under the guise of privilege, and Plaintiffs have sought to meet and confer with Meta to challenge those claims), making critical the testimony of key decisionmakers. Meta’s argument

⁷ Meta also argues that Messrs. Zuckerberg and Ah-Dahle were mentioned in the article, but they were not stated to have been a part of the potentially decisive “separate meeting.”

that “Mr. Cox appears in fewer than 100 of the over 20,000 Meta documents produced to date” only highlights Meta’s failure either to create or to produce relevant documents, leaving Plaintiffs little left to “exhaust” before getting directly from Mr. Cox the information that is otherwise unavailable.⁸

Meta’s argument that Plaintiffs may obtain similar information from two of Meta’s Directors of Product Management, Mike Clark and Chaya Nayak, falls flat given Mr. Cox’s central (and, as opposed to Clark and Nayak, leadership) role in an issue at the core of this case: data acquisition.

2. There Is No Reason to Limit Mr. Cox’s Deposition to a Mere Two Hours

Perhaps in recognition that its last apex gambit failed, Meta alternatively requests that the Court limit Mr. Cox’s deposition to a mere two hours—some five hours less than the default rule. Fed R. Civ. P. 30(d)(1). But Meta just repeats the same faulty arguments it makes in favor of blocking Mr. Cox’s deposition altogether. As already demonstrated, Mr. Cox has first-hand—and perhaps unique—knowledge, and Plaintiffs should not be cabined to a fleeting two hours to explore it, especially when much of that will surely be taken up by objections.

Meta contends that apex depositions must be time limited. But it depends on the needs of the case. In *In re Apple iPhone Antitrust Litig.*, Apple contended that its CEO, Tim Cook, should be deposed for only four hours, but the court concluded that “seven hours is the default rule” and “Plaintiffs have shown they can make good use of that time.” No. 11-CV-06714, 2021 WL 485709, at *3 (N.D. Cal. Jan. 26, 2021).⁹ Likewise, Meta has not shown that “the balance of relevance, proportionality, and burden warrant” limiting Mr. Cox’s deposition. *Garner v. Amazon.com, Inc.*, No. 21-CV-00750, 2023 WL 6038012, at *1 (W.D. Wash. Sept. 15, 2023). Here, Meta has not provided any reason why this particular deposition must be time limited. Notably, Meta notably does not claim that Mr. Cox’s job duties render him too busy for the deposition, which is a standard argument in apex motions.

⁸ Moreover, Meta has refused to make Mr. Cox a custodian, which Plaintiffs requested on October 20, such that his lack of appearance in documents so far produced is not indicative of much. This is particularly true as Plaintiffs learned for the first time during an October 16th meet-and-confer that Meta largely only searched and produced work email and other non-custodial company data sources like Workplace (Meta’s online collaborative software tool that includes instant messaging, and file sharing, and conferencing) and WhatsApp (which Meta owns) for its designated document custodians. See e.g., N.D. Cal. ESI Discovery Checklist, § III (discovery must be prioritized from systems like email, finance, HR); The Sedona Conference, The Sedona Principles, 19 Sedona Conf. J. 1, 103 (2018) (“sources of non-custodial relevant information” include “structured systems and databases, and other non-custodial sources such as collaboration tools, social media,” and any other “shared areas (such as public folders, discussion databases, and shared network folders) that are not regarded as belonging to any specific employee”). When Plaintiffs sought to inquire further into Meta’s searches, Meta claimed work product.

⁹ See also *Bos. Ret. Sys. v. Uber Techs., Inc.*, No. 19-CV-06361, 2023 WL 6132961, at *3 (N.D. Cal. Sept. 19, 2023) (permitting deposition of CEO for seven hours); *Jackson Fam. Wines, Inc. v. Zurich Am. Ins. Co.*, No. 22-CV-07842, 2024 WL 3325485, at *3 (N.D. Cal. July 8, 2024) (denying protective order seeking 5-hour limit); *Zimmerman v. Al Jazeera Am., LLC*, 329 F.R.D. 1, 7 (D.D.C. 2018) (denying protective order seeking 1.5-hour limit).

Plaintiffs have lived by the seven-hour rule without complaint, even though they, too, are busy individuals of significant stature. Meta deposed Sarah Silverman for 6 hours and 51 minutes, Christopher Golden for 6 hours and 54 minutes, David Henry Hwang for 6 hours and 39 minutes, Andrew Sean Greer for 6 hours and 37 minutes, Laura Lippman for 6 hours and 50 minutes, and Jacqueline Woodson for 6 hours and 58 minutes. There is no reason Mr. Cox—a witness with relevant, first-hand knowledge of facts critical to resolution of this case—should get preferential treatment.

III. Conclusion

For the foregoing reasons, Mr. Cox’s deposition should not be barred or restricted beyond the default rules.

B. Meta's Objections to Plaintiffs' Amended 30(b)(6) Deposition Notice / Motion for Protective Order

Meta's Position:

Introduction.

As an initial observation, Plaintiffs have turned what should have been a straightforward joint brief as to “deposition scheduling” into many pages of briefing on their gripes over document collection, custodians, and search terms – all governed by already passed deadlines in the Scheduling Order, governed by an ESI Order, and none appropriately presented to this Court in response to its Discovery Order, ECF No. 242. This has unfortunately detracted from the timely filing of this joint letter brief. In the little time provided, Meta has attempted to address Plaintiffs’ last minute, untimely and waived attacks as to ESI and document issues, but to the extent the Court engages with any of them, Meta asks for a reasonable opportunity to address any such issues with the Court.

Plaintiffs served an amended 30(b)(6) notice on October 8, 2024, spanning thirteen broad topics, many of which are overbroad and/or have little to no relevance to the claims and defenses in this case, including the core issue of fair use. Lauter Decl. Ex. B. Nevertheless, Meta agreed to present a witness on ten topics, with certain limitations, while objecting to the remaining three topics on grounds that they are overbroad and burdensome, concern irrelevant subject matter, would violate the parties’ stipulated ESI Order, and would improperly require Meta to provide witness testimony on information protected by attorney-client privilege and the work product doctrine.

Meta served responses and objections to Plaintiffs’ 30(b)(6) notice on October 28, 2024 (Lauter Decl., Ex. C), and the Parties met and conferred on October 29, 2024. Meta spent considerable time and effort narrowing its written objections and working with Plaintiffs to reach a compromise on topics that could be reasonably narrowed.¹⁰ As a result, the parties were able to reach agreement on most topics. Despite these efforts, the parties have been unable to reach agreement on Topics 5, 11, and 12, to which Meta objects in their entirety, and Topics 4 and 8, for which Meta has agreed to designate a witness as to portions of the topic, while maintaining its objections to other portions. Meta respectfully requests that the Court relieve it of any obligation to designate or prepare a witness on Topics 5, 11, 12 or the objectionable portions of Topics 4 and 8.¹¹

Legal Standard. Parties may “obtain discovery regarding any nonprivileged matter that is relevant to any party’s claim or defense and proportional to the needs of the case.” Fed. R. Civ. P. 26(b). “A party seeking to preclude a Rule 30(b)(6) deposition may apply to the court for a protective order under Federal Rule of Civil Procedure 26(c)(1).” *United States v. HVI Cat*

¹⁰ Plaintiffs’ histrionics about responses to correspondence are not pertinent to any of the issues. This case is fast moving and both sides are using their best efforts to keep up with the substantial correspondence being sent on a daily basis. As the record reflects, the parties were engaged and able to address all of the disputes several weeks before the 30(b)(6) depositions are scheduled to start.

¹¹ Plaintiffs’ suggestion in a footnote that Mike Clark’s deposition should be postponed is without merit. Plaintiffs’ have custodial documents from Mr. Clark, and their untimely letter on privilege log issues is no excuse for postponing his deposition. It has been difficult to set up the deposition calendar as it is and changing dates now will only risk pushing the schedule beyond the December 13 deadline.

Canyon, Inc., No. CV 11-5097 FMO (SSx), 2016 WL 11683593, at *5 (C.D. Cal. Oct. 26, 2016) (quoting *Bowers v. Mortg. Elec. Registration Sys., Inc.*, No. 10-4141-JTM, 2011 WL 6013092, at *3 (D. Kan. Dec. 2, 2011)). Under Rule 26(c)(1), the Court may, for good cause, issue an order to protect a party from “annoyance, embarrassment, oppression, or undue burden or expense” by “forbidding the disclosure or discovery.” Fed. R. Civ. P. 26(c)(1). Indeed, the Court must grant a protective order if “the discovery sought is unreasonably cumulative or duplicative, or can be obtained from some other source that is more convenient, less burdensome, or less expensive.” Fed. R. Civ. P. 26(b)(2)(C)(i).

A party propounding a 30(b)(6) deposition notice has an obligation to make the noticed topics “reasonably particularized” and craft matters of examination with “painstaking specificity.” See *HVI Cat Canyon, Inc.*, 2016 WL 11683593, at *3, *7. The topics for a 30(b)(6) deposition must also be relevant to the claims and defenses in the litigation. Fed. R. Civ. P. 26(b). Deposition topics lacking any metes and bounds are facially overbroad and unduly burdensome. See *Pres. Techs. LLC v. Mindgeek USA, Inc.*, No. 2:17-cv-08906-DOC-JPR, 2020 WL 10965163, at *5 (C.D. Cal. Oct. 19, 2020) (stating “deposition topic lacking all metes and bounds is facially overbroad and unduly burdensome”).

Argument. As addressed in detail below, Topics 5, 11 and 12 concern irrelevant, duplicative, and/or privileged subject matter, and are therefore not proper subjects of discovery. Additionally, Topics 4 and 8, as written, are overly broad, and not reasonably particularized such that Meta could adequately prepare a corporate witness to testify to them. Meta accordingly asks for an order relieving it of any obligation to prepare a witness on Topics 5, 11 and 12 and limiting the scope of required corporate testimony for Topics 4 and 8, as set forth below.

A. Topic 5 Improperly Seeks Testimony Regarding Irrelevant and Unspecified Foreign and Domestic Regulations

Topic 5 calls for the following testimony:

“Meta’s policies and procedures regarding its adherence to any laws or regulations, including the EU AI Act, that require transparency around the datasets used to train AI models, require authorization from rightsholders for the use of the text and data in training AI models, and require disclosure of the use of any copyrighted materials to train an AI model.”

Over the course of ten months, none of Plaintiffs’ interrogatories or document requests called for any discovery relating to that subject, and with good reason: this case is not about regulations around AI training data transparency. When asked about the relevance of Topic 5 during the parties’ meet and confer, Plaintiffs offered only conclusory statements of relevance, without any specific explanation of how or why. Lauter Decl. at ¶¶ 6–8. Notably Plaintiffs also were unable or unwilling to identify any specific U.S. laws or regulations on this topic as to which they were looking for testimony, referring vaguely to laws in California and Colorado. *Id.* ¶¶ 8–10. Based on Plaintiffs’ briefing below, they appear to be referencing AB 2013, a new California law that imposes certain transparency requirements around AI training data (and has nothing to do with copyright infringement or fair use), and the Colorado Artificial Intelligence Act (which

concerns high risk AI models, and also has nothing to do with copyright infringement or fair use). *See id.* Exs. D-E.

Rule 30(b)(6) “expressly requires that the party requesting the deposition ‘must describe with reasonable particularity the matters for examination.’” *See Memory Integrity, LLC v. Intel Corp.*, 308 F.R.D. 656, 661 (D. Or. 2015) (citing Fed. R. Civ. P. 30(b)(6)). District Courts within this Circuit note that “the party noticing a Rule 30(b)(6) deposition must take care to designate, with painstaking specificity, the particular subject areas to be covered.” *Mailhoit v. Home Depot U.S.A., Inc.*, No. CV 11-03892 DOC (SSx), 2012 WL 12884049, at *2 (C.D. Cal. Aug. 27, 2012) (ordering party to narrow its 30(b)(6) deposition topics where they were “impermissibly vague” and were not designated with “painstaking specificity”); *Izzo v. Wal-Mart Stores, Inc.*, No. 2:15-cv-01142-JAD-NJK, 2016 WL 409694, at *2 (D. Nev. Feb. 2, 2016) (internal quotation marks and citations omitted); *see also Tumbling v. Merced Irrigation Dist.*, No. 1:08cv1801 LJO DLB, 2009 WL 5064994, at *1 (E.D. Cal. Dec. 16, 2009). A topic that calls for testimony about “any laws or regulations” has no specificity at all, let alone the detail required by Rule 30(b)(6). Courts “have not hesitated to issue protective orders when corporations are asked to respond to overly broad or unfocused Rule 30(b)(6) deposition notices.” *HVI Cat Canyon, Inc.*, 2016 WL 11683593, at *8 (citation omitted).

In any event, Plaintiffs cannot explain how Meta’s compliance with any required transparency regulations is “relevant to any party’s claim or defense,” as required by the Federal Rules of Civil Procedure. Fed. R. Civ. P. 26(b). For example, the EU AI Act (the only law specifically mentioned in the Topic) is, to say the least, a complicated law that is not yet implemented, in relevant part, and governs various uses of AI in Europe. Lauter Decl., ¶¶ 12–13, Ex. F. This case does *not* concern implementation of EU Laws, alleged infringement in Europe, or compliance with foreign laws. Plaintiffs have never argued as much. Meta should not have to prepare a witness to discuss its understanding of the highly nuanced EU AI Act, much less the company’s future plans to implement procedures in response to that Act. That testimony is overly broad, irrelevant to the issues in this case, and would likely involve identifying somebody in Europe with either a legal background and/or whose knowledge is based on privileged communications. Whether Plaintiffs are seeking testimony concerning irrelevant EU regulations or irrelevant state statutes in the U.S., this Topic simply does not comport with Rule 26,

Plaintiffs’ justification for the relevance of this Topic, articulated below, is nonsensical. First, Plaintiffs now suggest that the Topic concerns “copyright infringement” but that term doesn’t appear anywhere within it. Second, Plaintiffs do not explain how Meta would even be subject to the EU AI Act with regard to challenged activities in the United States. Third, nothing in the article cited by plaintiffs indicates that Meta is refraining from releasing its models in the US or EU due to copyright concerns. To the contrary, public reporting indicates that the EU regulations Meta is concerned about are privacy-related. Lauter Decl. Ex. H. Furthermore, the authority cited by Plaintiffs is inapposite. The topic in *In re Hair Relaxer Mktg. Sales Pracs. & Prod. Liab. Litig.* No. 23-CV-00818, 2023 WL 8935006, at *3 (N.D. Ill. Dec. 27, 2023), concerned whether a defendant had notice of “safety risks” on the basis of foreign regulations; but what is “safe” for

consumers does not depend on local laws. What constitutes intellectual property infringement—or fair use—does.¹²

Plaintiffs also now suggest, for the first time, that certain laws implicate the industry definition of “open source,” which Plaintiffs claim is in dispute. This is not an issue that Plaintiffs have raised before now (i.e., within this letter brief) and Meta is unaware of any purported dispute over the definition of the term. In any case, whether Meta’s at-issue models meet any purported standard for “open source” has nothing to do with the issues the parties are litigating.

To the extent Plaintiffs are truly seeking testimony about “any laws or regulations,” it remains the case that Meta’s efforts to comply with transparency and disclosure requirements or mitigations for high risk AI models, if they exist, have no bearing on the issues in dispute. In addition, Meta is unaware of any laws “requir[ing] authorization from rightsholders for the use of the text and data in training AI models” nor have Plaintiffs identified any. Preparing a witness on unspecified laws is an impossible task and certainly one that is not proportionate to the needs of this litigation.

B. Topic 11 Improperly Seeks Discovery on Discovery

This Court previously declined to compel discovery in response to document requests that it perceived as discovery on discovery, explaining that the “Court has an obligation to keep discovery from metastasizing like this.” Discovery Order, ECF No. 153. Despite the Court’s guidance, Topic 11 is exclusively aimed at increasing the burden on Meta by seeking discovery on discovery that already occurred:

“Meta’s process for identifying and collecting documents in response to Plaintiffs’ RFPs, including data sources reviewed, non-custodial files reviewed, search terms used, date ranges applied, metadata collection and capabilities concerning what metadata may be collected, and Meta’s policies and practices relating to labeling internal documents “attorney-client privileged” or the equivalent.”

Discovery on discovery “is disfavored and, to be both relevant and proportional to the needs of the case, a party seeking it ‘must show a specific deficiency in the other party’s production.’” *Uschold v. Carriage Servs., Inc.*, No. 17CV04424JSWEDL, 2019 WL 8298261, at *4 (N.D. Cal. Jan. 22, 2019) (quoting *Brewer v. BNSF Ry. Co.*, No. CV-14-65-GF-BMM-JTJ, 2018 WL 882812, at *2 (D. Mont. Feb. 14, 2018)). See also *In re Google RTB Consumer Priv. Litig.*, No. 21-cv-02155-YGR (VKD), 2023 WL 5836816, at *6 (N.D. Cal. Sept. 8, 2023) (noting that “rather than targeting its inquiry to the particular log or data source at issue, plaintiffs ask for testimony about everything [defendant] did with respect to preservation of potentially responsive information for the entire case”); *Jensen v. BMW of N. Am., LLC*, 328 F.R.D. 557, 566 (S.D. Cal. 2019) (denying

¹² To the extent Plaintiffs attempt to tie the relevance of the EU AI Act to “willfulness,” that is nonsensical, for several reasons. First, as explained above, the EU Act requires compliance in EU countries, not the U.S., and does not address the fair use of training data. Fair use is a creature of U.S. law, and different standards may apply in the EU. Second, alleged knowledge that a download or copy was unauthorized (what Plaintiffs characterize as “willfulness”) is not supported by the case law, nor is it a factor in fair use. See 17 U.S.C. § 107.

motion to compel and explaining that “Plaintiff provides no particularized reason whatsoever for why such discovery should be ordered here, besides that this is ‘litigation.’”); *Ashcraft v. Experian Info. Sols., Inc.*, No. 2:16-cv-02978-JAD-NJK, 2018 WL 6171772, at *2 n.2 (D. Nev. Nov. 26, 2018) (“Discovery into another party’s discovery process is disfavored.”), *aff’d sub nom. Ashcraft v. Welk Resort Grp.*, No. 2:16-cv-02978-JAD-NJK, 2021 WL 3017512 (D. Nev. Mar. 24, 2021). Plaintiffs have not demonstrated any specific deficiency that would warrant deviating from the procedures set forth in the stipulated ESI Protocol.

Plaintiffs spend much of their response accusing Meta of “unilaterally” selecting custodians and search terms, *but that is what the parties negotiated and stipulated to*. On April 10, 2024, the Court entered the parties’ Joint Motion for Entry of Stipulated Order Re: Discovery of Electronically Stored Information. ECF No. 101. That 17-page ESI Order governs the procedures for selecting custodians (the producing party selects the “ten (10) custodians believed most likely to have discoverable ESI”) and exchanging information concerning the parties’ document collection efforts and reflects the parties’ agreement regarding the types of information that would be disclosed. *Id.* ¶ 6. Meta complied with the ESI Order regarding the selection of custodians, the identities of whom Plaintiffs have known since May (and which Plaintiffs have already made the subject of motion practice). The parties also agreed to disclose data sources, and Meta has complied with these obligations. The parties also agreed that “each Producing Party is best situated to determine the most appropriate method or methods for that Producing Party to search, collect, cull, and produce documents responsive to discovery” *Id.* ¶ 7. Plaintiffs’ complaint about search terms is ironic given Meta has offered a mutual exchange of search terms, which Plaintiffs have refused in favor of the unilateral approach they are seeking here.¹³ As to the propriety of testimony relating to “search terms,” many of the cases cited in Plaintiffs’ section below (received less than three hours before this filing) are ten years old; it does not appear that those cases employed an ESI Order, as is the case here. In any event, the ESI Order does not contemplate or permit 30(b)(6) testimony on these matters.

Plaintiffs purport to justify a deposition based on various alleged discovery disputes. None of the issues rise to the level of needing a deposition or are analogous to the cases Plaintiffs cite. In those cases, there was a specific issue that could not be resolved through other means, such as “failure to agree on search terms.” To the contrary, Plaintiffs’ arguments here prove Meta’s point - if Plaintiffs had questions or concerns about Meta’s discovery efforts, the ESI Order provided a mechanism for raising them. But to demand a deposition of a “discovery witness” in the hopes of manufacturing more discovery disputes or obtain discovery information it already has is improper and duplicative of the ESI Order protocols.

C. Topic 12 Improperly Seeks Corporate Testimony Regarding a 22-Page Legal Analysis Prepared by Meta Attorneys

Topic 12 is broadly directed at a 22-page legal analysis on fair use in the context of training large language models, i.e., the same legal issues the parties are litigating in this case:

¹³ In an October 21, 2024 email exchange, Meta represented the following to Plaintiffs: “As we stated, we are agreeable to engaging with Plaintiffs in a mutual exchange of search terms and hit counts as contemplated by the parties’ agreed-upon ESI order. (Doc No. 101, para. 7).”

“The factual bases and any other supporting analysis for Meta’s October 30, 2023 submission to the U.S. Copyright Office contending that its use of copyrighted material to train its LLMs is either non-infringing or otherwise constitutes ‘fair use.’”

This Topic is thus “an improper attempt to discover legal conclusions and/or protected information.” *See Google LLC v. Sonos, Inc.*, No. 20-cv-06754-WHA (DMR), 2022 WL 16554695, at *4 (N.D. Cal. Oct. 31, 2022) (citing *In re Indep. Serv. Orgs. Antitrust Litig.*, 168 F.R.D. 651, 654 (D. Kan. 1996) (stating that a party is not required to have counsel “prepare a witness to be able to testify on a given defense or counterclaim,” because such topics are “overbroad, burdensome, and a highly inefficient” and ordering that defendant not be required to produce any Rule 30(b)(6) witnesses)).

In *Lenz v. Universal Music Corp.*, this Court rejected a similar attempt to depose a 30(b)(6) witness on “the basis for any belief by Universal that Ms. Lenz’s video infringes the copyright in ‘Let’s Go Crazy,’ including without limitation *the basis for any belief by Universal that Ms. Lenz’s video is not a fair use of ‘Let’s Go Crazy.’*” No. C 07-03783 JF (PVT), 2010 WL 1610074, at *2–3 (N.D. Cal. Apr. 20, 2010). The court held that “[t]he facts that form those ‘beliefs’ are legal conclusions and an improper topic for a Rule 30(b)(6) deposition.” *Id.* at *3 (citing *3M Co. v. Kanbar*, No. C06-01225 JW (HRL), 2007 WL 1794936, at *2 (N.D. Cal. June 19, 2007) (finding that challenged topics “each inquire about support for the allegations that Defendant’s product infringes [plaintiff’s] marks” and holding that “in these circumstances, topics containing the terms ‘identical,’ ‘confusingly similar,’ and ‘blurring, dilution and tarnishment’ are, in effect, seeking legal conclusions that should not form the basis for 30(b)(6) deposition topics.”)). Topic 12 is even more obviously improper than the topic in *Lenz*, as it explicitly contemplates Meta’s “analysis” supporting the Copyright Office submission. Plaintiffs respond by mischaracterizing *Lenz*, ignoring that it directly addressed a topic concerning the “basis” of a copyright fair use claim.

Moreover, the topic does not identify what “factual bases” or “analysis” it seeks to cover from the submission, which encompasses a variety of discrete subjects ranging from the technical inner workings of large language models to economic markets for AI training data. Lauter Decl., Ex. G (Meta’s Comments in Response to U.S. Copyright Office Notice of Inquiry on Artificial Intelligence & Copyright). The Topic is thus inherently improper for lacking the required particularity. For this reason, too, the Topic should be stricken.

In addition, to the extent Plaintiffs are entitled to probe subject matters encompassed by the U.S. Copyright Office submission, they will certainly have the opportunity to ask about them in the 21 other fact depositions scheduled, via expert testimony, or in the context of other topics. For example, there are several fact witnesses capable of testifying about how Meta’s large language models work. There are accordingly other, better means (that do not run afoul of Rule 26’s proportionality requirement) to address any factual representations that Plaintiffs wish to test.

For all of these reasons, the Court should strike Topic 12.

D. Topics 4 and 8 are Overbroad and Should be Construed in Accordance with Meta’s Proposals

Plaintiffs’ position that Meta prepare a witness to testify to Topics 4 and 8, as written, is impracticable, unreasonable, and disproportionate to the needs of the case. These Topics lack the necessary “painstaking specificity” required of 30(b)(6) topics. *Izzo*, 2016 WL 409694, at *2. *Pres. Techs.*, 2020 WL 10965163, at *5 (issuing protective order as to facially overbroad and unduly burdensome topic). A broad 30(b)(6) topic “render[s] unworkable the obligation of the responding party to make a conscientious, good-faith effort to designate knowledgeable persons . . . and to prepare them to fully and unequivocally answer questions about the designated subject matter.” See *Apple Inc. v. Samsung Elecs. Co.*, No. C 11-1846 LHK (PSG), 2012 WL 1511901, *2 (N.D. Cal. Jan. 27, 2012) (cleaned up).

Topic 4 – This Topic calls for:

Meta’s knowledge that its LLM training datasets included copyrighted material, including any discussions or deliberations regarding the same; any discussions, deliberations, or efforts to mitigate potential copyright infringement relating to Meta’s LLMs; and any steps Meta has taken to train or otherwise program its LLMs not to output infringing material.

Meta has agreed to designate a witness to testify regarding “Meta’s knowledge of whether certain datasets used to [train the Meta Language Models contain material that is protected by copyright under U.S. law and Meta’s efforts to reduce the possibility, if any, that the models can reproduce verbatim content contained within training data.” Meta has declined only to provide a designee concerning “discussions or deliberations” about these matters.

Plaintiffs have numerous depositions remaining, wherein Plaintiffs will have an opportunity to ask each witness about the datasets and any non-privileged discussions¹⁴ and deliberations concerning them. Plaintiffs have also served document requests related to (if not overlapping with) this subject matter and, to the extent that relevant, responsive, non-privileged documents reflecting such discussions or deliberations can be located following a reasonable search, they have been or will be produced. As a result, to the extent such discussions are relevant, Plaintiffs have an avenue to obtain discovery on them from the most reliable source – documents and the witnesses themselves. But to require Meta to either (a) prepare a witness to testify about all of those same discussions (if they occurred) or (b) embark on an impossible task of investigating discussions that occurred anywhere within the company is unduly burdensome and should not be allowed. “[I]t is simply impractical to expect Rule 30(b)(6) witnesses to know the intimate details of everything.” *Fed. Ins. Co. v. Delta Mech. Contractors, LLC*, No. CA 11-048ML, 2013 WL 1343528, at *4 (D.R.I. Apr. 2, 2013); *Dealer Computer Servs., Inc. v. Curry*, No. 12 Civ. 3457(JMF)(JLC), 2013 WL 499520, at *2 (S.D.N.Y. Feb. 7, 2013) (“A [30(b)(6)] deposition is not a quiz, nor is it the most practical way to obtain [all types of] information.”).

¹⁴ Plaintiffs improperly raise arguments concerning whether privilege applies to certain communications in their response below. The issue of what particular communications are privileged is not ripe or before the Court. The parties are separately addressing privilege issues, with Meta sending a thorough response in response to Plaintiffs’ allegations, which are neither well-taken nor timely.

Topic 8 – This Topic calls for:

Meta’s communications including discussions, deliberations, and negotiations concerning each and every dataset regardless of the source, type, or potential use that may or could have been used in training generative AI products (whether or not related only to Meta products) or for use in any other Meta AI product, other than generative AI, including but not limited to, all such discussions referenced in the April 6, 2024 *New York Times* article, “How Tech Giants Cut Corners to Harvest Data for A.I.” or the September 25, 2024 *The Verge* article and interview, “Why Mark Zuckerberg thinks AR glasses will replace your phone.” Some examples that are obvious may include actual or potential licensing deals related to Meta’s AI assistant, or actual or potential licensing deals related to Meta’s AI instant video generator, regardless of whether they resulted in actual licensing deals. This topic includes all communications related to these discussions, deliberations, negotiations and or attempts by Meta or other generative AI companies that may have been discussed by Meta that did not result in an executed licensing agreement with a third party.

Meta has agreed to designate a witness to testify regarding “Meta’s efforts to acquire text data to train the Meta Language Models, including licensing efforts, if any,” which is what Meta assumed Plaintiffs were seeking with this Topic. During the parties’ meet and confer, however, Meta was surprised to learn that Plaintiffs construe this Topic as broadly as it is written, to encompass every “discussion[], deliberation[], and negotiation[] concerning each and every dataset regardless of the source, type, or potential use that may or could have been used in training generative AI products” of any kind, including image, video, and audio generation models, none of which are at issue in this dispute. Meta’s efforts to acquire data for *other*, undefined, “generative AI products”—a topic that is not at issue in Plaintiffs’ Complaint and, heretofore, has not been the subject of discovery—has no bearing on Plaintiffs’ copyright infringement claims or Meta’s fair use defense.

Below, to support the overbreadth of Topic 8, Plaintiffs cite an article discussing deals with celebrities for “AI chatbots.” This only underscores that the Topic is overbroad and would impose on Meta an undue burden to prepare a witness on wholly irrelevant subject matter. Personas or individual likenesses licensed for AI chatbot products are obviously dissimilar to the issues in this case, which is about alleged copyright infringement of Plaintiffs’ fictional books in the training of Meta’s LLMs – not about an AI chatbot’s ability to sound just like John Cena. The instant case is about copyright law and training, not publicity rights associated with the creation of outputs. Thus, testimony or documents on purported deals with celebrities for their *voices* – and the specific terms of those deals – are not relevant here, nor would it be proportionate for Meta to have to prepare a corporate witness to recite those deal terms.

In addition, and as with Topic 4, it would be unduly burdensome, if not impossible, to require Meta to prepare a witness on a subject matter of this breadth and depth. *Every* discussion? *Every* deliberation? About *every* dataset that *may* or *could* have been used for *any* generative AI product of any modality? A Rule 30(b)(6) deposition is an inappropriate vehicle for obtaining the information plaintiffs seek and is unduly burdensome in light of Meta’s agreement to provide testimony on a narrower, more reasonable construction of Topic 8. *See Google*, 2022 WL 16554695, at *4 (citing 168 F.R.D. at 654 (granting a protective order where the Rule 30(b)(6) request was “overbroad, burdensome, and a highly inefficient method through which to obtain

otherwise discoverable information’’)). Plaintiffs respond by suggesting that Meta has “near limitless resources,” and that this somehow obviates the impossibility, to say nothing of the lack of proportionality, of preparing a corporate witness to testify regarding the broad subject matter encompassed by the Topic as written. Even if the quip about Meta’s resources were true, it is beside the point.

Plaintiffs’ Position on 30(b)(6) Deposition:

1. Introduction

Meta’s efforts to substantially curtail Plaintiffs’ reasonable, proportionate, and targeted 30(b)(6) discovery topics are prejudicial to Plaintiffs and consistent with its repeated efforts to limit and unilaterally define the scope of this critically important case. Its outcome may determine whether content creators are appropriately compensated when large corporate interests steal their intellectual property without permission and profit from it.¹⁵

Although Meta seeks to limit the case to what it claims is the “core issue” of fair use, fair use is simply a defense—one of many Meta apparently intends to assert—to Plaintiffs’ claims. Plaintiffs are entitled to *all* discovery relating to their claims, not just discovery cabined to Meta’s preferred affirmative defense. It is understandable that Meta may want to focus on “fair use.” That helps it justify its profit-maximizing business decision to engage in widespread criminal and fraudulent activity by knowingly downloading and using massive amounts of copyrighted works without licensing them and then scrubbing copyright information from those works to cover up their acts. But fair use is not the “core issue” in this case—Meta’s infringement of Plaintiffs’ works is.

Finally, even if fair use were the only issue in this case, Meta cannot limit relevant and proportional discovery into it. Contrary to Meta’s unsupported contention, the topics at issue - those where Meta is refusing to produce a witness, or to limit the topic - relate, albeit non-exclusively, to fair use. One topic relates to a gating issue of all discovery (Topic 11), another involves Meta’s business efforts to claim its misconduct *was fair use* (Topic 12), a third involves the commercial purposes of the uses (and the extant and developing licensing market for training data) (Topic 8), and others go directly to willfulness and Meta’s knowledge (Topics 4, 5).

While Plaintiffs have made every effort to work cooperatively with Meta, many of those efforts have been rebuffed or ignored altogether by Meta, as detailed further below in the Factual Background section. Ultimately, Meta’s complete refusal to designate anyone at Meta to testify regarding three noticed topics—and to seek to materially curtail testimony on two others—is inconsistent with the general broad scope of discovery permitted under the Federal Rules and case law, and is further misplaced for the reasons described below.

¹⁵ Just yesterday, Meta’s announced its latest earnings, which saw “Meta profits surge 35% on AI-driven progress.” According to Meta’s CEO, Mark Zuckerberg, “We had a good quarter driven by AI progress across our apps and business. We also have strong momentum with Meta AI, Llama adoption, and AI-powered glasses.” See Ex. H, “Meta profits surge 35% as an AI spending spree pays off,” *available at* <https://qz.com/meta-platforms-q3-2024-earnings-report-1851685123>.

2. Factual Background

As this Court is aware, prior to the involvement of new counsel and the enlargement of the discovery schedule, Plaintiffs served an eleven-topic 30(b)(6) deposition notice on Meta on September 16, 2024. Meta sought to limit 30(b)(6) deposition testimony to 12 hours. On October 1, 2024, this Court ordered that “Plaintiffs may take a total of 16 hours of deposition testimony on their Rule 30(b)(6) deposition notice, separate and apart from their 30(b)(1) depositions. How to allocate those 16 hours between Rule 30(b)(6) designees *is for Plaintiffs to decide.*” Dkt. 196 (emphasis added).

Two business days after Judge Chabbria granted a two-month extension of the initial case management deadline for the close of fact discovery, Plaintiffs promptly served an amended 30(b)(6) deposition notice on Meta on October 8, 2024. The notice covers many of the same topics as the original notice, and adds in a handful of additional discrete topics. *See* Plaintiffs’ Ex. I. Contemporaneous with serving the amended notice, Plaintiffs’ counsel offered to meet and confer with Meta at its earliest convenience regarding any questions or objections Meta may have to the amended notice. Plaintiffs also stated: “We are happy to discuss scheduling at your earliest convenience given Judge Hixson’s order regarding the amount of time Plaintiffs will have.” *See id.* Notably, despite the subsequent expansion of the discovery schedule and changes to the notice, Plaintiffs have not sought additional time beyond the 16 hours this Court previously ordered.

On October 15, 2024, this Court ordered the parties to immediately facilitate the scheduling of all remaining depositions. That same day, Plaintiffs reached out to Meta to request that Meta identify the designees for the noticed 30(b)(6) topics so the parties could work to schedule those depositions. Ex. J. Meta did not respond. Plaintiffs’ counsel followed up via email on Thursday, October 17, once again asking that Meta identify deponents for the noticed topics. *Id.* Meta did not respond to that inquiry either.

During a meet and confer on October 16, 2024, as memorialized in an email from Plaintiffs’ counsel, Meta expressly declined to state whether it would object to any of the amended topics served on October 8 and refused to meet and confer on that issue. Ex. K. On Monday, October 21, 2024, Meta responded: “As we told you, we were evaluating your topics and identifying the witnesses that Meta plans to designate.” *Id.* Meta further represented: “Meta will meet and confer with Plaintiffs’ this week about the 30(b)(6) topics and any objections.” *Id.* Shortly after receipt of that email, Plaintiffs replied: “Thank you, please let us know when Meta is available to meet and confer this week. Also, please let us know if Meta will identify its designated 30(b)(6) witnesses (and by topic) this week. Absent Meta’s agreement to provide this information, we remain at an impasse over this issue.” *Id.*

Despite Meta’s representation that it would meet and confer during the week of October 21-25, Meta did not respond to Plaintiffs’ request to set a time, requiring a further follow-up from Plaintiffs on the morning of Thursday, October 24 requesting “some proposed times so we can move this issue forward promptly.” Ex. L. Meta replied nearly twelve hours later offering availability only on the following Tuesday, October 29. *Id.* In light of Meta’s unexplained delay, Plaintiffs replied on Friday asking that “in order to ensure we keep moving things forward expeditiously, we will need Meta to provide its deponent list for each topic with availability and objections by noon pacific on

Monday so we can assess in advance of the M&C and so we can provide an update for the depo status report due Tuesday to Judge Hixson.” *Id.* Meta once again did not respond to Plaintiffs’ request, so on Sunday, October 27, Plaintiffs confirmed the Tuesday meet and confer “since that is the earliest time [Meta] have made available,” but reiterated Plaintiffs’ request to receive Meta’s list of “deponents and availability for each topic, as well as Meta’s objections” by noon Pacific on Monday, October 28 “in order to ensure the M&C proceeds efficiently.” *Id.* Yet again, Meta did not respond, so Plaintiffs reached out again on Monday afternoon:

Following up on the several emails we have sent on this topic already, to which Plaintiffs have received no response, will Meta please advise on the witnesses it will be designating for Plaintiffs’ 30(b)(6) topics and their availability for deposition, as well as any objections Meta may have to the noticed topics? We have been seeking this information for weeks now (the amended notice was nearly three weeks ago), and with numerous depositions already being calendared, Meta’s refusal to provide this information is prejudicing Plaintiffs’ ability to seek discovery.

Id.

Meta did not respond to this email. Instead, Meta served its objections and responses on the evening of Monday, October 28 (*i.e.*, the night before the scheduled meet and confer). Ex. M.

The parties met and conferred on Tuesday, October 29, during which Meta once again indicated it was not prepared to identify its designees for the 30(b)(6) deposition notice. Meta first provided designees for some, but not all, 30(b)(6) deposition topics at 4pm Pacific yesterday, October 30, 2024.¹⁶ Ex. N. Meta also provided its final position regarding objections and proposed limitations to Plaintiffs’ 30(b)(6) deposition topics at the same time yesterday, leaving Plaintiffs only a matter of hours to research and draft responses to Meta’s requests for a protective order as to five different topics. *Id.*

3. Legal Standard

Courts in the Northern District of California generally allow for a broad scope of 30(b)(6) testimony. Depositions taken pursuant to Fed. R. Civ. P. 30(b)(6) may properly seek any evidence which may lead to the discovery of admissible evidence. *Vietnam Veterans of Am. v. Cent. Intel. Agency*, 2010 WL 11730757, at *9 (N.D. Cal. Nov. 12, 2010); *see also Detoy v. City and County of San Francisco*, 196 F.R.D. 362, 366-67 (N.D. Cal. 2000) (holding that the scope of 30(b)(6) deposition is determined solely by relevance under Rule 26).

The rule of “reasonable construction” applies to Rule 30 (b)(6) topics. A 30(b)(6) topic is not objectionable if the topic may be reasonably construed to seek relevant information capable

¹⁶ As a result, Plaintiffs learned only yesterday that Meta has designated Mike Clark, who is scheduled to be deposed in two weeks, and for whom Meta has logged more privileged documents than it has produced, for five key 30(b)(6) topics. As discussed elsewhere in this submission, Plaintiffs respectfully request that the Court order Meta to provide additional dates for his deposition between December 3 and December 10 so Plaintiffs can receive Meta’s response to its privilege letter and then brief the remaining disputed issues before this Court before Mr. Clark’s deposition.

of being addressed in a deposition. *See Luna v. Universal City Studios LLC*, 2015 WL 13655668, at *6 (C.D. Cal. Feb. 10, 2015); *Westmoreland v. Regents of the Univ. of California*, 2019 WL 932220, at *3 (E.D. Cal. Feb. 26, 2019).¹⁷ Accordingly, if topics or definitions may be reasonably construed to seek relevant information in a particular deposition, such topics or definitions are permissible. *United States ex rel. Poehling v. Unitedhealth Grp., Inc.*, No. CV 16-08697 FMO, 2022 WL 19913651, at *1 (C.D. Cal. Nov. 30, 2022).

Further, the party seeking to limit discovery has the burden of proving “good cause,” which requires a showing “that specific prejudice or harm will result” if the protective order is not granted. *In re Roman Catholic Archbishop of Portland*, 661 F.3d 417, 424 (9th Cir. 2011).

4. Argument

Meta does not argue, and cannot show, that it will suffer some “specific prejudice or harm” from having to prepare one or more witnesses to testify regarding Topics 5, 11, and 12, as discussed further below. Rather, Meta generally claims that those topics “concern irrelevant, duplicative, and/or privileged subject matter.” Purported objections to relevance or duplication are not valid bases to refuse to provide testimony and instead seek a protective order. *See Tri-State Hosp. Supply Corp. v. United States*, 226 F.R.D. 118, 125–26 (D.D.C. 2005) (“Nevertheless, I am aware of no principle of law that precludes a party from pursuing during a deposition a [30(b)(6)] topic about which it has already received information via other discovery devices. By its very nature, the discovery process entails asking witnesses questions about matters that have been the subject of other discovery.”)); *see also id.* noting that “[p]arties may obtain discovery regarding any matter, not privileged, that is relevant to the claim.”); *Lopez v. City of La Habra Heights*, 2023 WL 4680796, at *6 (C.D. Cal. June 1, 2023) (holding that “Plaintiff was not permitted to refuse to answer questions at the deposition” where Plaintiff made several objections as to relevance); *Rabin v. Google LLC*, 702 F. Supp. 3d 880, 881 (N.D. Cal. Nov. 20, 2023) (“[T]he party seeking protection bears the burden of showing specific prejudice or harm will result if no protective order is granted.”).

Similarly, rather than justify a total bar on testimony, Meta’s privilege objections can be addressed as privilege issues are during the normal course of deposition testimony. To the extent Meta believes that any particular question calls for privileged information, it of course can and should instruct the designee not to answer and Plaintiffs can seek court intervention should it believe such an instruction is inappropriate.

With respect to Topics 4 and 8, the testimony sought relates to critical issues in the case, including (a) Meta’s knowledge and willful disregard that it was stealing copyrighted material (which goes directly to fair use factor one) and (b) the effect of Meta’s theft on the potential market for copyrighted works (fair use factor four). Given that Meta is already preparing witnesses on those topics, the incremental burden associated with preparing those witnesses to testify to two narrow additional areas (for Topic 4, “discussions and deliberations” at Meta regarding Meta’s

¹⁷ Meta’s reliance on *United States v. HVI Cat Canyon, Inc.*, No. CV 11-5097 FMO (SSX), 2016 WL 11683593, at *5 (C.D. Cal. Oct. 26, 2016), is inapposite because there the deposition notice at issue covered “at least sixty-one” topics, whereas here, Plaintiffs have limited their notice to 13 targeted topics.

knowledge that it was using copyrighted material for AI training, and for Topic 8, efforts to acquire or license non-textual AI training data, such as images, audio, or video) is de minimis.

Plaintiffs therefore request that the Court deny Meta’s request for an order relieving it of any obligation to prepare a witness on Topics 5, 11 and 12 and limiting the scope of required corporate testimony for Topics 4 and 8.

Topic 5 – Meta May Not Refuse to Prepare a Witness Regarding Relevant Testimony That Bears Directly on Meta’s Willful Copyright Infringement

Topic 5 calls for the following testimony:

“Meta’s policies and procedures regarding its adherence to any laws or regulations, including the EU AI Act, that require transparency around the datasets used to train AI models, require authorization from rightsholders for the use of the text and data in training AI models, and require disclosure of the use of any copyrighted materials to train an AI model.”

Meta’s claim that Topic 5 does not seek relevant information is belied both by the broad scope of permissible discovery, including for 30(b)(6) depositions as well as by analogous case law. *See, e.g., Wachuku v. JetBlue Airways Corp.*, 2021 WL 4497157, at *7 (C.D. Cal. Apr. 29, 2021) (finding policies and procedures in place at the time of the incident at issue relevant and discoverable). Specifically, any steps Meta has taken to abide by regulations relating to AI models, including regulations seeking to prevent copyright infringement are relevant to, among other issues, the nature and willfulness of Meta’s infringement, which is a critical part of fair use factor one. Notably, because the EU has passed regulations intended to counteract copyright infringement by large corporations, Meta has refused to release its latest Llama models in the EU, specifically blaming the “the unpredictable nature of the European regulatory environment” according to a Meta spokesperson. *See* Ex. O, “Meta won’t release its multimodal Llama AI model in the EU,” *available at* <https://www.theverge.com/2024/7/18/24201041/meta-multimodal-llama-ai-model-launch-eu-regulations>.¹⁸

Meta’s decision to curtail its AI products in response to certain regulations is compelling evidence that sheds light on Meta’s knowledge and notice of its own allegedly criminal, fraudulent conduct in misappropriating copyrighted works. *See In re Hair Relaxer Mktg. Sales Pracs. & Prod. Liab. Litig.*, No. 23-CV-00818, 2023 WL 8935006, at *3 (N.D. Ill. Dec. 27, 2023) (internal citations omitted).

Here, information related to products developed and sold by L’Oréal S.A. outside the United States can shed light on L’Oréal U.S.A.’s knowledge and notice of safety risks. Foreign regulatory communications are thus relevant ‘to the extent they

¹⁸ Although Meta purports to be ignorant regarding certain provisions of the EU AI act, numerous analyses and practice guides describing the EU AI Act make clear that it will require companies such as Meta to get “express consent from rightsholders” for use in the development and training of AI models, including by “identifying and respecting any rights reservations expressed by copyright holders.” *See, e.g.,* Ex. R, Chambers and Partners, “EU AI Act: The Essential Guide to Copyright Compliance for General-Purpose AI Models,” (Sept 2, 2024), *available at* <https://chambers.com/articles/eu-ai-act-the-essential-guide-to-copyright-compliance-for-general-purpose-ai-models>

contain information about what [L'Oréal U.S.A.] knew about the alleged risks associated with their [hair relaxer] products, when they knew about those alleged risks, and whether those alleged risks were communicated' to regulators, distributors, and consumers.

See also Brunston v. Bayer Healthcare Pharms., Inc., No. EDCV131904FMODETBX, 2014 WL 12587032, at *3 (C.D. Cal. May 16, 2014) ("Plaintiffs are requesting information and documents regarding the use of Avelox in other countries, including information regarding labeling, safety, and adverse events . . . Plaintiff contends that these documents are relevant to showing what defendants knew regarding the safety of Avelox. Plaintiffs further argue that defendants used foreign clinical trials to support the approval of Avelox in the United States and defendants' safety department or central pharmacovigilance unit is located in Germany. . . Here, the Court concludes that, generally, plaintiffs have sufficiently demonstrated the relevance of foreign documents and information.").

Moreover, contrary to Meta's inaccurate representation of what transpired during an October 28, 2024 meet and confer, in addition to the EU AI Act, Plaintiffs' counsel specifically identified laws passed by Colorado and California that are also within the scope of the noticed topic. The law passed in California, of which Meta is undoubtedly aware because it actively lobbied against a companion measure, requires companies such as Meta to provide transparency around the training data used to train their large language models. *See* Ex. P, California A.B. 2013, "Artificial Intelligence Training Data Transparency." Indeed, just this week, public reports emerged that Meta has taken "issue with a new definition of open source AI that would require model creators to detail their sources of training data, among other rules, to meet the standard." Ex. Q, "Meta, OSI tussle over definition of open source AI," *available at* <https://www.axios.com/2024/10/29/meta-osi-definition-open-source-ai-llama>. Given that Meta has both publicly and in this litigation tried to claim that its models are "open source" such that fair use should apply to its infringement, Meta's policies and procedures relating to the issue of training data transparency are highly relevant to the disputed factual question of whether Meta's LLMs satisfy the industry standards to be characterized as "open source."

Given that Meta initially identified its training data for Llama 1, but then stopped doing so after it began relying even more heavily on copyrighted material, testimony regarding Meta's policies and procedures to abide by a California law and other regulations that will require such transparency is relevant to the nature and willfulness of Meta's infringement.

Topic 11 – Meta May Not Refuse to Prepare a Witness Regarding Important Issues Concerning How Meta Has Complied with Its Discovery Obligations

Topic 11 calls for the following testimony:

Meta's process for identifying and collecting documents in response to Plaintiffs' RFPs, including data sources reviewed, non-custodial files reviewed, search terms used, date ranges applied, metadata collection and capabilities concerning what metadata may be collected, and Meta's policies and practices relating to labeling internal documents "attorney-client privileged" or the equivalent.

With respect to Topic 11, the documents produced thus far raise serious questions about the efficacy of Meta’s discovery searches. Despite the Parties’ efforts to meet and confer, questions remain—and those questions require answers, necessitating Rule 30(b)(6) testimony on this Topic. The testimony will be a vehicle, and likely the *most efficient* vehicle, for Plaintiffs to obtain necessary information for use during upcoming witness depositions and in connection with extant and pending discovery dispute briefing (including over the role of Meta’s in-house counsel in the approval of materials for “training,” and other issues relating to Meta’s privilege claims, including its withholding of several hundred documents and communications relating to business issues) *See* Ex. CC, October 23, 2024 Letter from Plaintiffs to Meta raising issues with Meta’s privilege assertions.

Simply because Meta has searched the custodial files of ten of its self-selected custodians with search terms it unilaterally selected (and which Plaintiffs still have not seen), does not mean that its search was reasonable. *See Beganovic v. Tyson Fresh Meats, Inc.*, No. 22-CV-2052-LTS-KEM, 2023 WL 9503379, at *2 (N.D. Iowa, Mar 12, 2023) (“A party may not decline to produce ESI simply because it was not specifically requested or because search terms and custodians were not identified.”). Courts in and out of the Northern District of California permit discovery into—and expect cooperation regarding—the opposing party’s effectuation of its discovery obligations. For example, in *United Ass’n of Journeyman & Apprentices of the Plumbing & Pipe Fitting Indus., Underground Util./landscape Loc. Union No. 355 v. Maniglia Landscape, Inc.*, 2019 WL 7877821, at *2 (N.D. Cal. July 25, 2019), the defendant stated it could not locate responsive documents, but refused to share information about its search process or search terms. Magistrate Judge Beeler denied a motion for a protective order and required ongoing conferral, noting that “[p]arties exchange this sort of information as part of the general discovery meet-and-confer process without the need for formal ‘discovery-on-discovery’ requests.” *Id.* at 2-3 (citing *Apple, Inc. v. Samsung Elecs. Co. Ltd.*, 2013 WL 1942163, at *1, *4 (N.D. Cal. May 9, 2013) (ordering third-party subpoena recipient to produce its search terms and custodians after requesting party asked for this information “not as part of a formal Request for Production nor as an Interrogatory but rather as a request following a meet-and-confer”); *see also Baranco v. Ford Motor Co.*, 2018 WL 9869540, at *1 (N.D. Cal. Apr. 10, 2018) (citing N.D. Cal. Guidelines for Discovery of ESI, 1.02, 2.02-2.03 for the proposition that it is “well-established that, when search terms are used in ESI discovery, the parties should cooperate to select reasonable search terms and custodians.”); *De Abadia-Peixoto v. U.S. Dep’t of Homeland Sec.*, 2013 U.S. Dist. LEXIS 120368, at *10 (N.D. Cal. Aug. 23, 2013) (ordering defendant to disclose search parameters and to meet and confer regarding their sufficiency). Even beyond search terms, courts have routinely held that noncustodial sources of ESI should be searched in conjunction with custodial sources. *E.g., Raine Grp. v. Reign Cap., LLC*, 2022 WL 538336, at *2 (S.D.N.Y. Feb. 22, 2022) (“Counsel for both parties must consult with their respective clients to understand which custodians *and locations* are likely to have relevant information whether or not responsive to its adversary’s discovery requests.”) (emphasis added).

Furthermore, designating 30(b)(6) witnesses specifically regarding a party’s search methodology is commonly accepted in the Northern District of California. For example, following parties’ failure to agree on search terms for accessible files, while facing an approaching deadline to complete focused discovery, the district judge in *Westly v. Oclaro, Inc.* granted a 30(b)(6) witness on several matters, including “steps taken to preserve, retrieve, collect and gather potentially relevant electronic data.” *Westly v. Oclaro, Inc.*, No. 11-CV-02448 EMC NC, 2013

WL 1365910, at *3, Dkt. No. 129-2 at 3 (N.D. Cal. Apr. 3, 2013); *see also Campbell v. Facebook Inc.*, 310 F.R.D. 439 (N.D. Cal. 2015) (“While Facebook may have provided detailed responses to Plaintiffs’ interrogatories and submitted the declarations of Facebook employees describing how its private message processing works, Plaintiffs have a right to verify and explore the information Facebook provided through a 30(b)(6) deposition.”); *FormFactor, Inc v. Micro-Probe, Inc.*, No. C-10-03095 PJH JCS, 2012 WL 1575093 (N.D. Cal. May 3, 2012) (“The March 2 order required Plaintiff to ‘sit for a 30(b)(6) deposition only as to topic # 30 regarding storage systems and searches undertaken to respond to Micro–Probe’s document requests.’”).

This practice extends throughout Ninth Circuit district courts. In *Puckett v. Country of Sacramento*, a magistrate judge ordered defendants to produce a witness to address, among other ESI topics, searches conducted to identify responsive documents after defendants “repeatedly stated that various documents never existed or were not maintained.” No. 2:22-CV-0350 KJM DB, 2024 WL 418187, at *1, *4 (E.D. Cal. Feb. 5, 2024), *reconsideration denied*, No. 2:22-CV-00350-KJM-DB, 2024 WL 1462390 (E.D. Cal. Apr. 4, 2024). The court noted the distinct value of 30(b)(6) depositions where a party has been unable to fully comply with discovery requests. *Id.* at *3; *see also M & F Fishing, Inc. v. Certain Underwriters at Lloyds*, No. 06CV0934-DMS (BLM), 2007 WL 9706491 (S.D. Cal. Apr. 13, 2007) (permitting 30(b)(6) deposition of document custodian where plaintiff had not produced any documents); *MetroPCS v. A2Z Connection, LLC*, No. 215CV01412JADDJA, 2020 WL 127550 (D. Nev. Jan. 10, 2020) (granting 30(b)(6) witness on discovery practices because “Plaintiff is entitled to know what categories of ESI Defendants preserved and collected and how a reasonable search for responsive documents was performed with sufficient specificity to demonstrate due diligence”); *Anstead v. Virginia Mason Med. Ctr.*, No. 221CV00447JCCJRC, 2022 WL 1641425, *5 (W.D. Wash. May 24, 2022) (denying defendant’s motion for protective order against plaintiff’s 30(b)(6) request for “the scope, nature, and method of search” where defendants lost relevant evidence).

Those cases and scores of others and this District’s guidelines tout the benefits of cooperation. Moreover, they underscore that parties are entitled to basic information about search parameters. Thus, while this Court would not need to find that Meta’s ESI productions were deficient to order Meta to produce a witness about at least some of the sub-topics in Topic 11, Plaintiffs submit Meta’s document production, and related information sharing, has been deficient. This is therefore a situation where “discovery on discovery” is warranted. *See Jensen v. BMW of N. Am., LLC*, 328 F.R.D. 557, 566 (S.D. Cal. 2019) (cited by Meta).

Here, among other things, the insufficiency and deficiency of Meta’s production is reflected by the fact(s) that Meta:

- (a) has refused to exchange search terms and hits counts as required by the ESI Order, insofar as it has failed to respond to multiple requests for a date to exchange them. *See* Ex. T.
- (b) has refused to identify potentially relevant data sources and witnesses, and even when it has produced information, it has done so tardily, and in a manner that prejudices Plaintiffs. For example, when Meta identified its data sources searched, Plaintiffs promptly noticed the data sources did not include text messages, WhatsApp messages and data from Meta’s Manifold and requested those sources be searched and produced. Ex. U. Meta

ignored those requests. Instead, Meta only produced a smattering of WhatsApp messages hours before the critical deposition of their VP of AI Ahmad Al-Dahle. Those messages, however, indicate that other witnesses (including other custodians and deponents) also routinely use WhatsApp to conduct business and those communications have yet to be produced. Ex. V.

(c) has refused to add new custodians even when Plaintiffs demonstrated a reasonable and proportionate basis to do so, and indeed, only added custodians upon court intervention;

(d) has exacerbated this problem by failing to search repositories and document sources outside of its self-selected group of custodians. Specifically, Plaintiffs learned for the first time during an October 16, 2024 meet and confer that Meta largely only searched and produced work email and other non-custodial company data sources like Workplace (Meta's online collaborative software tool that includes instant messaging, and file sharing, and conferencing) and WhatsApp (which Meta owns) for its designated document custodians. *See, e.g.*, Exs. W & X. N.D. Cal. ESI Discovery Checklist, § III (discovery must be prioritized from systems like email, finance, HR); The Sedona Conference, The Sedona Principles, 19 Sedona Conf. J. 1, 103 (2018) ("sources of non-custodial relevant information" include "structured systems and databases, and other non-custodial sources such as collaboration tools, social media," and any other "shared areas (such as public folders, discussion databases, and shared network folders) that are not regarded as belonging to any specific employee"). When Plaintiffs sought to inquire further into Meta's searches, Meta claimed work product. And with respect to relevant repositories, documentary and deposition testimony have revealed [REDACTED]

Ex. Y.

(e) appears to have shrouded swathes of its AI business operations in privilege, including designating many documents as "Attorney Client Privilege" "A/C privilege" or its equivalent. *See e.g.*, Exs. Z; AA; BB; *See also* Ex. CC.

In fact, each sub-topic in the notice corresponds to one of these particularized concerns. Given Meta's delays and conduct in withholding necessary information, and the limited time left in discovery, a fair and efficient outcome is a deposition on the relevant topic of Meta's document sources.

Finally, Meta's arguments against designating a witness are misplaced. *First*, Meta asserts that the deposition would cause discovery to expand. This response is puzzling. This would only be true if it became apparent based on sworn testimony that Meta had shielded relevant and proportional discovery. And if that were true, there would be more serious implications than merely the expansion of discovery.

Second, Meta argues that the ESI order does not anticipate a witness on ESI discovery. From Plaintiffs' perspective, however, it is Meta's disregard of the ESI order that explains in part why this deposition is appropriate to begin with. Meta also argues that Plaintiffs should have raised concerns with compliance earlier, but this ignores that Plaintiffs in fact have been doing exactly that. *E.g.*, Exs DD; EE [8/22/24 Benon Ltr to J. Lauter]; [9/19/24 Stameshkin Email Chain re Data Sources].

Third, Meta acknowledges, as it must, that even so-called “discovery on discovery” can be appropriate. The cases it cites confirm this, and are factually distinguishable, generally because the moving party did not attempt to set forth deficiencies, was merely anticipating future deficiencies, and/or because (unlike here) the non-moving party offered alternative and reasonable means to provide the sought information. *See, e.g. In re Google RTB Consumer Priv. Litig.*, 2023 WL 5836816, at *6 (N.D. Cal. Sept. 8, 2023) (in the context of a concern over potential spoliation, noting that defendant Google had offered to provide verified interrogatory responses about the preservation of data). It is plain here that Meta’s discovery is deficient, and Topic 11 is therefore appropriate.

Topic 12 - Meta May Not Refuse to Prepare a Witness Concerning Its Public Comment to the Copyright Office Based on an Erroneous, Wholesale Assertion of Privilege

Topic 12 calls for the following testimony:

The factual bases and any other supporting analysis for Meta’s October 30, 2023 submission to the U.S. Copyright Office contending that its use of copyrighted material to train its LLMs is either non-infringing or otherwise constitutes “fair use.”

Meta objects to Topic 12, which requests testimony on “the factual bases and any other supporting analysis for Meta’s October 30, 2023 submission to the U.S. Copyright Office” (a submission referred to here as the “Public Comment”) on the basis that the Topic impermissibly seeks to “discover legal conclusions and/or protected information”; that the Topic “lacks the required particularity” due to the purported breadth of the Public Comment; and because other fact witnesses “capable of testifying about how Meta’s large language models work” and expert witnesses will also testify. None of these arguments is a valid basis for allowing Meta not to prepare a 30(b)(6) deponent on Topic 12.

First, in seeking to limit 30(b)(6) topics it is Meta’s burden to prove “good cause,” meaning “specific prejudice or harm will result” without a protective order. *Phillips ex rel. Ests. of Byrd v. Gen. Motors Corp.*, 307 F.3d 1206, 1210 (9th Cir. 2002) (“the party seeking protection bears the burden of showing specific prejudice or harm will result if no protective order is granted”). Yet Meta does not engage at all with the content of the Public Comment, except by *mischaracterizing* it, in entirely conclusory fashion, as a “15-page legal analysis on fair use.” In fact, the Public Comment was Meta’s publicly filed and publicly available report responding to an invitation by the U.S. Copyright Office’s (U.S.C.O.) seeking “*factual information* and views on a number of copyright issues raised by recent advances in generative AI.” Ex. FF_[12-A that Reed sent at midnight]_ (emphasis added). The U.S.C.O. sought input from industry participants and the public to “advise Congress; inform its regulatory work; and offer information and resources to the public, courts, and other government entities considering these issues.”

The Public Comment contains sections concerning fair use, but also sections that are clearly relevant and that can in no way be considered privileged—*e.g.*, sections titled “AI Models and Their Function” (which itself has a sub-section titled “How Large Language Models ‘Learn’”) and “Meta’s Role in the AI Landscape.” *See* Ex. GG [12-B that Reed sent at midnight]_ at 2-10.

Indeed, Meta concedes just a few paragraphs later in its briefing that the Public Comment “encompasses a variety of discrete subjects ranging from the technical inner workings of large language models to economic markets for AI training data”—topics that are completely disconnected from Plaintiff seeking any “legal conclusions” from Meta.

In any event, there is no rule categorically prohibiting Plaintiffs from asking Meta about affirmative defenses in any deposition. Litigants are permitted to take discovery on matters relevant to another party’s defense, including in 30(b)(6) depositions. *Smith v. City of Bakersfield*, 2024 WL 71687, at *3 (E.D. Cal. Jan. 5, 2024) (citing *Stevens v. Brigham Young University*, 420 F. Supp.3d 1114, 1123 (D. Id. 2019)) (permitting Rule 30(b)(6) topic deemed relevant to rebutting defense). Meta, naturally, has the right *at the deposition* to object in good faith to questions to protect privileged communications, and to instruct the deponent not to answer on that basis. Meta may not, however, use its threadbare, erroneous showing in this letter briefing to bar such testimony. Meta’s broad-brush, false characterization of the entirety of the Public Comment as exclusively providing “legal analysis on fair use” fails to show any “specific prejudice or harm” in requiring Meta to prepare and designate a witness to testify about the Public Comment.

Second, Meta’s cases are easily distinguishable. Plaintiffs seek testimony concerning the “factual bases and any other supporting analysis” for a report Meta officially submitted to the government in a fully public document. Most sections, if not every section, of the report are relevant to the subject matter of this action, and only some portions of the report address copyright or fair use. Those circumstances are nothing like 30(b)(6) topics seeking a contractual party’s “understanding of” provisions in the contract at issue between litigants (as in *Google LLC v. Sonos, Inc.*, 2022 WL 16554695, at *4 (N.D. Cal. Oct. 31, 2022)) or “the basis for any belief” by an entertainment company that a musician’s video infringes the company’s copyright (as in *Lenz v. Universal Music Corp.*, 2010 WL 1610074, at *2–3 (N.D. Cal. Apr. 20, 2010)). Plaintiffs may ask a huge variety of relevant questions about the Public Comment that have nothing to do with any of Meta’s “understanding” or “beliefs” about anything privileged. *See Luna v. Universal City Studios LLC*, 2015 WL 13655668, at *6 (C.D. Cal. Feb. 10, 2015) (topic not objectionable if it may be reasonably construed to seek relevant information capable of being addressed in a deposition); *United States ex rel. Poehling v. Unitedhealth Grp., Inc.*, 2022 WL 19913651, at *1 (C.D. Cal. Nov. 30, 2022) (if topics or definitions may be reasonably construed to seek relevant information in a particular deposition, such topics or definitions are permissible). Wholesale exclusion of the Public Comment as a topic for preparation and testimony by Meta representatives is unsupportable and unwarranted.

Finally, Meta provides no authority supporting either of its contentions that (1) Topic 12 is “inherently improper for lacking the required particularity” or (2) Meta need not prepare a witness for Topic 12 because there are purportedly “better means to address any factual representations” in the Public Comment “that Plaintiffs wish to test,” by asking other fact witnesses. As to (1), it is not surprising that Meta could not provide support for applying the “reasonable particularity” standard in a limiting way. To the contrary, this Court has held that the “reasonable particularity” requirement “cannot be used to limit what is asked of the designated witness at a deposition,” and “[t]he 30(b)(6) notice establishes the minimum about which the witness must be prepared to testify, not the maximum.” *UniRAM Tech., Inc. v. Monolithic Sys. Tech., Inc.*, 2007 WL 915225, at *2 (N.D. Cal. Mar. 23, 2007). Even if such an application of the

rule were allowable, there is nothing unfair about asking Meta to provide testimony about its own official communication prepared and submitted just *one year ago*.

Similarly, Meta cannot avoid a 30(b)(6) topic on the basis that there is a purportedly “better means” to prospectively pursue testimony from others. Rather, the rule is that Meta must “not only produce such number of persons as will satisfy the request, but more importantly, prepare them so that they may give complete, knowledgeable and binding answers on behalf of the corporation.” *Guifu Li v. A Perfect Day Franchise, Inc.*, 2011 WL 3895118, at *2 (N.D. Cal. Aug. 29, 2011). Notably, Meta has not argued (because it cannot) that the content of any part of the Public Comment is irrelevant. And Meta also has not argued (because it cannot) that preparing deponents to testify on the subjects in the Public Comment is overly burdensome. Meta concedes that its own employees—the deponents in the other scheduled depositions—“are capable of testifying about how Meta’s large language models work,” the subject of extensive sections of the Public Comment. The Court should require Meta to prepare for and testify on Topic 12.

Topic 4 – Meta’s Proposed Limitation Is Inappropriate and Would Deprive Plaintiffs of Relevant Evidence

This Topic calls for:

Meta’s knowledge that its LLM training datasets included copyrighted material, including any discussions or deliberations regarding the same; any discussions, deliberations, or efforts to mitigate potential copyright infringement relating to Meta’s LLMs; and any steps Meta has taken to train or otherwise program its LLMs not to output infringing material.

Meta has refused to designate a witness to testify concerning “discussions or deliberations” that go to the very heart of this case relating to Meta’s knowledge “that its LLM training datasets included copyrighted material.” Far from being overbroad, the requested testimony is narrowly tailored to a very discrete issue, which distinguishes the noticed topic from the cases Meta cites. Nor is there a lack of specificity here – Meta is well aware what the topic requests, and objects primarily on the basis of a purported burden. But as Plaintiffs’ counsel stated during the parties’ meet and confer, Meta’s obligation to prepare a deponent to testify on this topic is reasonably cabined by what is required of a 30(b)(6) witness under the federal rules, including the case law cited by Meta.

Meta presumably does not want to offer testimony on this narrow issue because such testimony will reveal, as reflected in a *New York Times* article regarding some of these discussions, that Meta made a conscious, deliberate business decision to illegally use copyrighted material in order to stay competitive in an AI arms race. *See* Ex. HH, How Tech Giants Cut Corners to Harvest Data for A.I., *available at* <https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html> (reporting that in March and April 2023, some Meta employees “met nearly daily to tackle the problem” of getting “more data,” including discussions “about copyrighted data with senior executives, including Chris Cox, Meta’s chief product officer,” where “no one in that meeting considered the ethics of using people’s creative works”).

Should the Court be inclined to narrow the scope of the “discussions or deliberations” at issue, pending the parties’ briefing and a decision on Meta’s privilege claims, Plaintiffs would be amenable to limiting the “discussions or deliberations” to the March and April 2023 time period wherein Meta employees met near daily to discuss these issues.

To the extent Meta may contend that testimony about these “discussions or deliberations” may implicate attorney-client privilege, “[a]n in-house counsel’s advice regarding business matters is not protected by attorney client privilege.” *Dolby Lab’s Licensing Corp. v. Adobe Inc.*, 402 F. Supp. 3d 855, 873 (N.D. Cal. 2019). Furthermore, “[w]hen a communication may relate to both legal and business advice, the proponent of the privilege must make a ‘clear showing’ that the ‘primary purpose’ of the communication was securing legal advice.” *Id.* at 873 (citing *United States v. ChevronTexaco Corp.*, 241 F. Supp. 2d 1065, 1076 (N.D. Cal. 2002)). Testimony addressing copyrighted data to train Llama models is a business activity because it primarily relates to operational, strategic, or technical aspects of the business rather than seeking legal advice on legal risks associated with using specific datasets. The “burden” is on Meta to “clearly show that the primary purpose” of the testimony is to secure legal advice, which Meta has not done. *Id.*

Moreover, even if Meta had provided or can provide sufficient information to show that that the primary purpose of this testimony was to give or receive legal advice, as opposed to business matters, this testimony nonetheless must be provided under the crime-fraud exception. This District has recognized the exception in the copyright infringement context. *See, e.g., Netbula, LLC v. Chordiant Software, Inc.*, No. 08-cv-0019-JW (HRL), 2009 WL 3919495, at *2 (N.D. Cal. Nov. 18, 2009).¹⁹

Topic 8 - Meta’s Proposed Limitation Is Inappropriate and Would Deprive Plaintiffs of Relevant Evidence

This Topic calls for:

Meta’s communications including discussions, deliberations, and negotiations concerning each and every dataset regardless of the source, type, or potential use that may or could have been used in training generative AI products (whether or not related only to Meta products) or for use in any other Meta AI product, other than generative AI, including but not limited to, all such discussions referenced in the April 6, 2024 New York Times article, “How Tech Giants Cut Corners to Harvest Data for A.I.” or the September 25, 2024 The Verge article and interview, “Why Mark Zuckerberg thinks AR glasses will replace your phone.” Some examples that are obvious may include actual or potential licensing deals related to Meta’s AI assistant, or actual or potential licensing deals related to Meta’s AI instant video generator, regardless of whether they resulted in actual licensing deals. This topic includes all communications related to these discussions, deliberations, negotiations and or attempts by Meta or other generative AI companies that may have been discussed by Meta that did not result in an executed licensing agreement with a third party.

¹⁹ Plaintiffs have raised the crime-fraud privilege exception to Meta in connection with challenging Meta’s overbroad, unsupported privilege assertions regarding documents, and Plaintiffs anticipate this issue will only be resolved through Court intervention. *See* Ex. CC, Oct. 23, 2024 Meet & Confer Ltr. from Pltfs’ Counsel to Meta Counsel re: Deficiencies in Meta Priv. Logs.

The Court should reject Meta’s efforts to exclude from Topic 8 highly relevant testimony regarding efforts Meta has made to license non-textual training data for its AI large language models or for use in any other Meta AI products, which are all built on the Llama models.

If Meta has sought to, or in fact has, licensed non-textual AI training data, that is relevant to several key issues in this case, including fair use factors one and four. Specifically, any actual or contemplated licensing deals are highly relevant to the fair use analysis that looks to the effect on the market for the infringed works. This is doubly true given that Meta has tried to claim, including through a public submission to the U.S. Copyright Office, that there is no real market for licensing AI training data. *See* Ex. II at pp. 16-17 (“Put simply, there is no legitimate market through which copyright holders can be compensated for the use of statistical facts extracted from their works”). Plaintiffs are entitled to test Meta’s assertion regarding the lack of a “legitimate market” for AI training data through the requested testimony.²⁰

Once again, Meta’s complaints about the scope of the topic are easily addressed by Meta adhering to the well-understood obligations imposed by Rule 30(b)(6). Moreover, requiring Meta to offer testimony regarding image, audio, or video training data that it may have acquired or sought to acquire will not add an undue burden on Meta, a company with near limitless resources.

Nor can Meta credibly claim that preparation for this topic will require some type of encyclopedic knowledge. Based on discovery reviewed by Plaintiffs and public reporting, Meta has engaged in highly-targeted efforts to acquire training data for its AI models and products, such that any designee could educate themselves adequately. For example, as was recently widely-reported, in an effort to “help its AI chatbot . . . keep pace with competitors’ products,” Meta signed deals with multiple celebrities “to train [its] chatbot to replicate their voices”. Ex. KK, “Meta is bringing the voices of Judi Dench, John Cena, and Keegan-Michael Key to its AI chatbot,” *available at* <https://www.cnn.com/2024/09/25/tech/meta-ai-celebrities-chatbots/index.html>. Testimony regarding the terms and conditions of this licensing deal for AI training data directly implicates Meta’s affirmative defense of fair use.

²⁰ Meta’s self-serving assertion has recently been refuted by its own announcement of a multi-year licensing deal with Reuters that “will allow Meta to use *Reuters* content for its chatbot responses.” Ex. LL, “Meta signs its first big AI deal for news,” *available at* <https://www.theverge.com/2024/10/25/24279259/meta-reuters-ai-chatbot-deal-news-licensing-media> (reporting that “*Reuters* will be compensated for its content appearing in Meta’s AI chatbot”).

C. Plaintiffs' Request RE: Meta's Production of Custodial Files

Plaintiffs' Position:

With nearly all depositions now scheduled, Plaintiffs have repeatedly raised issues with Meta about its extremely limited production of custodial files from some of Meta's original ten self-selected custodians, such as Joelle Pineau. On Monday, October 21, 2024, Plaintiffs' counsel emailed Meta's counsel about this issue. [Pritt email to P. Morton et al 4:46pm on 10/21] Meta did not respond to that email regarding this issue. Plaintiffs' counsel followed up with emails on Friday, October 25, and again on Monday, October 28, 2024. Meta did not respond to either follow-up email regarding this issue.

Meta's consistent refusal to respond to Plaintiffs' inquiries is striking given that Ms. Pineau is scheduled for deposition next week, on November 6. Despite being one of Meta's original custodians, Meta has only produced 355 documents from Ms. Pineau's custodial file, compared to thousands of custodial files from certain of Meta's other document custodians. Ms. Pineau appears on only **56** emails and almost no Workplace chats (Workplace was Meta's online collaborative software tool that includes instant messaging, file sharing, and video and audio conferencing), let alone any WhatsApp messages (Meta owns WhatsApp). Given Meta's refusal to even discuss this issue, Plaintiffs respectfully request that the Court order Meta to immediately re-review its collection and search for Ms. Pineau's responsive documents and communications, and to make a supplemental production by November 4.

Plaintiffs have identified a similar issue with the custodial file of Mike Clark, who is scheduled for deposition in two weeks and who Meta, for the first time yesterday afternoon, identified as the 30(b)(6) deponent on five topics, including some that cover the most critical issues in this case, such as Meta's copying of datasets containing the works at issue in this case and use of those works to train Meta's Llama models. Despite being a custodian, Meta has only produced **199** documents from Mr. Clark's custodial file. Not only is this a shockingly low number of files for a custodian in and of itself, let alone Meta's Director of Product Management – Generative AI, but it is even more astonishing when compared to the number of documents and communications from Mr. Clark's custodial files withheld under the guise of attorney-client privilege. Mr. Clark is listed as the custodian of 155 out of 467 documents on Meta's non-email privilege log and an additional 90 documents on Meta's non-email log. Mr. Clark is a software engineer, not an attorney, yet Meta has withheld more of his custodial documents than it has produced. Plaintiffs have raised substantial concerns with Meta's privilege claims in a meet-and-confer letter on October 23, 2024, including that it has withheld massive amounts of information that appear to concern business issues or, at a minimum, cannot satisfy the primary purpose test, or otherwise must be produced under the crime/fraud exception. See Ex. CC. Meta has not responded substantively to that letter despite Plaintiffs' request. Given that Meta has now designated Mr. Clark as its corporate representative on several critical topics in this case yet has produce little from his custodial files and withheld an even larger amount of his files under privilege claims that Plaintiffs will challenge, Plaintiffs respectfully request that the Court (1) order Meta to immediately re-review its collection and search of Mr. Clark's responsive documents and communications, including those on its privilege logs, and to make a supplemental production by November 4, and (2) order Meta to provide additional dates for his deposition between December 3 and December 10 so Plaintiffs can receive Meta's response to its privilege letter and then brief the remaining disputed issues before this Court before Mr. Clark's deposition.

Relatedly, to date, Plaintiffs have not received *any* custodial documents from the five documents custodians the Court added on October 4, despite their upcoming depositions. *See* Dkt. 212 (ordering Meta to add Mark Zuckerberg, Eleonora Presani, Amanda Kallet, Alex Boesenberg, and Nikolay Bashlykov as custodians). Although Meta stated yesterday for the first time that they are working on these productions, this near-month delay is inconsistent with Meta's earlier representation with respect to Mr. Zuckerberg that they could produce his custodial file within two weeks. Indeed, Meta's representations on these issues have been constantly shifting. *Compare* 9/25 Hr'g Tr. at 17:11-15 (representing to the Court that it takes "a minimum of two weeks to produce documents from even one custodian") *with* Dkt. 190, ¶ 2 (stating in a declaration that it "can take up to one week or longer for a single ESI custodian"). Even based on its own shifting representations, Meta should have produced all responsive documents in all five of these custodians' files, including Mr. Zuckerberg, by now. Plaintiffs have abided by the Court's earlier guidance that production of documents from new custodians should not slow the pace of scheduling (or taking) depositions, and Plaintiffs have diligently worked toward that end for each of Meta witnesses. However, it appears that Meta's strategy is to run out the clock so as to deprive Plaintiffs of any reasonable opportunity to examine deponents with these additional custodial documents. Accordingly, Plaintiffs respectfully request that the Court order Meta to produce all custodial documents from the five custodians it ordered Meta to add a month ago by no later than November 4, 2024.

Finally, Plaintiffs further request that Meta be ordered to include in any subsequent collection and production any responsive WhatsApp messages, of which Meta has produced only a handful in its entire production. Given that Meta owns WhatsApp, Meta should be obligated to treat collection from WhatsApp akin to collection from email and other company systems. *See, e.g.,* N.D. Cal. ESI Discovery Checklist, § III (discovery must be prioritized from systems like email, finance, HR).

In response to these issues, Meta largely fails to address the merits of the issues Plaintiffs raise. Instead, Meta doubles down on its continued attempt to "run out the clock" and avoid fulfilling even basic discovery obligations that have come to light recently.

First, Meta argues that the subject of deponent custodial files is not appropriately raised here. Yet, the very text it quotes shows the topic of deponent's document productions is germane, as the Court asks the parties to address "every single issue that is outstanding concerning depositions." ECF No. 243).

Second, Meta argues that the issue is not timely because of language in the stipulated deadlines concerning *written* discovery before the Court. Those deadlines—proposed by Plaintiffs—not only concerned previously-served written discovery requests (and thus are irrelevant here), but it was predicated upon a shared understanding that such interim deadlines would encourage the parties to work expeditiously and in good faith to continue to meet and confer, produce ESI search terms, conduct reasonable searches in response to discovery requests, and so on; that way, the extended discovery schedule could be productive in the manner the Court envisioned. Clearly Meta has not been acting in accordance with that understanding, which Plaintiffs outlined in their briefing to the Court. ECF No. 244.

Third, Meta, again not addressing the merits, argues that Plaintiffs notified Meta too late that they would raise the issue before the Court. Not so. Plaintiffs raised in both of their discovery deficiency meet and confer letters on October 9, 2024 (Ex. NN at pp. 2-3, 7; Ex. OO at pp. 2, 4-5)

and again repeatedly during the parties' meet and confer on October 16. In any event, Meta's argument just shows that Plaintiffs have acted diligently to meet the Courts' deadlines—Plaintiffs continued to try to negotiate a resolution to the issue in good faith and only after Meta's continued delays and final refusal to even offer a compromise did Plaintiffs decide they had no choice but to raise this issue with the Court. Meta essentially asks this Court to penalize Plaintiffs for trying to avoid bringing a dispute before it. The Court should decline to do so.

Fourth, when Meta does finally turn to the "merits," it argues over minutiae, disputing for example whether the correct document count for Mr. Clarke is 255 or 196. Notably, Meta does not and cannot provide any cognizable evidence suggesting that its productions were complete (and its collections of ESI thorough) because Meta has refused to produce its search terms and hit counts (another item Plaintiffs have had to raise before the Court because Meta continually rebuffed Plaintiffs' efforts until it finally returned its response: that this issue too was now untimely). *See* ECF No. 244. Indeed, Meta does not dispute that it failed to search for and produce responsive documents from non-custodial data sources like its company email servers and other non-custodial company sources like Workplace (Meta's instant messaging, and file sharing, and conferencing software) and WhatsApp (another of Meta's IM systems), but rather only searched those systems for their very limited number of custodians.

Finally, Meta also argues that it needs more time to complete ESI collections because of its prioritization of Mr. Zuckerberg's data collection. Meta is a massive company with nearly unlimited resources. It simply beggars belief for Meta to suggest that it has made so little progress on ESI collection that it cannot promptly produce custodial documents for remaining and additional custodians. *See, e.g., Epic Games, Inc. v. Apple Inc.*, Case No. 20-cv-05640-YGR (TSH) (ND Cal. Sept. 27, 2024) (Hixon, J.) ("Apple is one of the largest companies in the world, with nearly infinite resources available to it. If Apple really wanted to, could it collect and review 1.3 million documents [in a tight time frame]? Yes, of course it could. If Apple really wanted to, with all of the resources available to it, it could probably review that many documents in a weekend."). Further, the Court should not credit Meta's assertion that "it can take over eight weeks to complete the process of producing custodial documents." Rule 34 requires production of documents "no later than the time for inspection specified in the request or another reasonable time specified in the response," which usually is the same time that a party must respond to the requests—30 days. The same is true under the California Discovery Act. *See* 2031.030, 2031.280. Meta cites nothing—no ESI articles or the like—that remotely supports Meta's professed inability to collection and produce documents in a timely fashion. Indeed, there is none, as Plaintiffs' vendors routinely collect custodial files in a single day, which can then be quickly processed and reviewed.

Meta's Position:

The Court should reject Plaintiffs' attempt to raise ESI custodial issues in the context of this brief for several reasons.

First, Plaintiffs' claimed issues with Meta's custodial document production are not issues related to the "scheduling of . . . depositions" (ECF No. 243). The Court's Order directing this letter briefing specifically directed the parties to address "every single issue that is outstanding concerning depositions...including all other issues that need to be resolved to finalize the scheduling of all remaining depositions." *Id.* The sufficiency of Meta's custodial document

productions has nothing to do with the scheduling of depositions, particularly in view of this Court's clear directive that the Plaintiffs should "expeditiously proceed with depositions with the documents they already have and don't wait until they get more documents." (ECF No. 231 at 3). Thus, Plaintiffs' demands should be rejected on this basis alone.

Second, Plaintiffs have failed to timely raise, and thus have waived, their purported issues with Meta's custodial productions. As they admit, they first raised this issue after the October 16, 2024 deadline for meeting and conferring on "Existing Written Discovery" (ECF Nos. 227, 238), when Plaintiffs buried the request in email chains that were principally about Meta's efforts to schedule depositions. (Lauter Decl., Ex I). If this issue was truly significant, Plaintiffs could have and should have raised it in advance of the stipulated October 16 meet and confer deadline, especially given that Meta produced 237 of Mike Clark's 256 custodial documents and 334 of Joelle Pineau's 342 custodial documents by August 30, 2024.²¹ Plaintiffs have offered no excuse for waiting so long to raise any concerns about the purported size of these custodians' document productions, and there is none. Plaintiffs also did not raise this (or any) issue in letter briefing to this Court before the stipulated, Court-ordered October 23, 2024 deadline for Letter Briefs re Existing Written Discovery (ECF Nos. 227, 238). Plaintiffs have effectively admitted their failure to comply with this deadline for Letter Briefs re Existing Written Discovery by seeking to extend that deadline (ECF No. 244) in a motion filed with Judge Chhabria that Meta will oppose on November 4. In short, Plaintiffs have waived any purported complaints about the sufficiency of Meta's custodial document productions, and unless the District Court grants Plaintiffs' pending motion to extend the deadline, this Court should not allow Plaintiffs to circumvent that deadline.

Third, Plaintiffs first notified Meta that they would move on these non-scheduling related document issues on the afternoon of October 30 – less than one day before today's noon filing deadline. If Plaintiffs had questions about the adequacy of prior custodial productions – and setting aside the aforementioned deadlines that they ignored – they should have outlined those issues for Meta, met and conferred (for example, on the nearly 3-hour meet and confer that took place on October 16), and provided Meta an adequate opportunity to respond. They did not. Forcing Meta to dispel Plaintiffs' factual misrepresentations in less than 20 hours is not the way to create a proper record. Moreover, any disputes about the sufficiency of custodial ESI production are addressed by the parties' agreed upon ESI Order. ECF No. 101, Ex. A, para. 7.

Nonetheless, Meta provides the Court with additional context as to why Plaintiffs' complaints about Meta's custodial productions are without merit. To begin, Plaintiffs' brief argues, incorrectly, that Meta has "refused" to respond to inquiries regarding Ms. Pineau's documents. Meta has diligently responded to a myriad of demands and requests from Plaintiffs' various counsel on a near-daily basis and alongside deposition scheduling, an ongoing effort to produce additional documents (including ESI ordered by the Court), and other briefing.

Plaintiffs also complain that Meta "only produced 355 documents" from Ms. Pineau's custodial files and "199" from Mr. Clark's custodial files. But the fact that the number of responsive documents is not larger is not evidence of any alleged "deficiencies"; it is simply the result of (1) the Requests for Production served by Plaintiffs and (2) the small volume of these

²¹ The small number of documents produced for Mr. Clark and Ms. Pineau after August 30 were the result of documents that were initially withheld but ultimately produced as Meta updated its privilege logs.

custodians' correspondence on issues responsive to the Requests for Production. Additionally, Plaintiffs served new discovery requests on October 9 and October 18, to which responses are not yet due. Subject to Meta's forthcoming objections and responses, Meta will produce any additional responsive custodial documents for the ten original custodians and the five additional custodians ordered by the Court as promptly as reasonably possible.

Additionally, Plaintiffs' representation regarding the number of documents produced for Mr. Clark is incorrect. Meta has produced approximately 256 custodial documents for Mr. Clark (not 199, as Plaintiffs assert). While it is true that a substantial number of Mr. Clark's documents are partially or completely withheld on the basis of attorney-client privilege, that is a product of Mr. Clark's role²², which is to work closely with the legal team on a number of issues related to Meta's Llama models beyond the issues of training data at issue in this case. Moreover, whether Mr. Clark may testify on certain 30(b)(6) topics does not correlate to his possession of documents, as corporate witnesses are often prepared to testify regardless of their possession of documents about a given topic. And certainly, none of Plaintiffs' speculative arguments justify *postponing* Mr. Clark's deposition – which has already been scheduled. As to Plaintiffs' assertions regarding Meta's privilege log, Meta is addressing Plaintiffs' meritless and untimely 16 page letter in separate correspondence.²³

Regarding the additional five ESI document custodians ordered by the Court on October 4, Plaintiffs never raised concerns about the pace of that production until yesterday afternoon, and did not reveal their demand that custodial documents be produced by November 4 until this letter brief. Indeed, Meta has been hard at work collecting, searching and reviewing those materials for production. Plaintiffs' brief is also wrong to suggest that Meta has "shifted" positions on the time it takes to complete those productions. As Meta explained in its prior briefing on custodial document production, it can take over eight weeks to complete the process of producing custodial documents. (ECF No. 190-8)²⁴ Meta certainly cannot complete multiple ESI custodial reviews by November 4, a few days from now, or even in a matter of several weeks, given the multiple

²² Mr. Clark is not currently working in a capacity as a "software engineer" as Plaintiffs initially asserted in their brief. Rather, he works in product management and in a role that requires him to work closely with in-house counsel, and given the number of legal issues that these products touch on, it follows that he would have a number of privileged discussions with the Meta legal team.

²³ As Plaintiffs note, they first sent a letter regarding Meta's Amended Privilege Logs on October 23, which was untimely in light of the stipulated, Court-ordered deadlines for meeting and conferring regarding existing written discovery (October 16) and filing letter briefs on any dispute regarding existing written discovery (October 23). *See* ECF Nos. 227, 239. Even though all of the issues in Plaintiffs' October 23 letter are untimely and therefore waived, Meta is substantively responding to that letter to make clear the meritlessness of Plaintiffs' complaints.

²⁴ As Meta already addressed in a prior brief, it never said that it "could produce [a custodian's] file within two weeks." ECF No. 190 at 7-8. Meta previously explained that "two weeks" was the absolute minimum time period if there was a *single* custodian and it diverted every resource at the company and its e-discovery vendor to the project. *Id.* Meta's prior briefing explained in detail the time-consuming process ESI discovery has taken, including that for 10 ESI custodians it took over 8 weeks to collect, review, and produce those documents. ECF No. 190 at 7-8; ECF No. 190-8. Plaintiffs also misquote Dkt. 190-8 Paragraph 2, which indicated that the *first step of collecting* and processing documents would take at least one week if done to the exclusion of other discovery tasks.

steps and volume involved. *Id.* The unreasonable timeframes for production demanded by Plaintiffs for the first time in this letter brief are simply not possible – as Meta explained the last time Plaintiffs demanded ESI productions within a short timeframe. ECF No. 190 at 7-8; ECF No. 195 at 2 (this Court noting that granting motion for more custodians “would be adding several weeks of additional document productions”).

Moreover, at the demand of Plaintiffs, Meta has specifically prioritized review and production of documents from Mr. Zuckerberg, which was a condition of the parties’ negotiated agreement regarding the timing and parameters of his deposition. Given Plaintiffs’ demands, Meta has been focusing its efforts on substantially completing production of Mr. Zuckerberg’s custodial documents by November 1 under the parties’ agreement governing his deposition, which is six weeks away and the final scheduled deposition in this case. Meta has prioritized production to meet the Plaintiffs’ demands and is deploying substantial resources to finish the remaining custodians as promptly as it reasonably can within the time periods it previously represented. Meta is doing this while preparing for and attending to over 20 depositions of its witnesses and numerous new discovery requests served after the Court extended the discovery cut-off.

Finally, Plaintiffs argue that because “Meta owns WhatsApp, Meta should be obligated to treat all” WhatsApp messages as “akin” to company email. This argument is legally baseless. It is also another “Existing Written Discovery” issue that Plaintiffs waived and which is inappropriate for resolution on less than 20 hours notice in a letter brief regarding deposition scheduling issues. Nonetheless, like all other custodial data sources, Meta has undertaken collections from WhatsApp messages when a custodian utilizes that application for work purposes, which is and was a small number of the overall custodians.

D. Plaintiffs' Request RE: Testimony from Former Meta Employees***Plaintiffs' Position:***

Plaintiffs have sought for weeks to procure deposition testimony from certain former employees of Meta, including Susan Zhang, Guillaume Lample, Edouard Grave, and Aurelian Rodriguez. The latter three reside in France and were critical in Meta's development of its LLMs, including the downloading and use of the copyrighted works at issue in this case. Plaintiffs have repeatedly asked Meta to provide the last known contact information for these witnesses so Plaintiffs can attempt to obtain documents and deposition testimony from them through the appropriate procedures if Meta will not make them available. To date, Meta has not responded. Notably, Mr. Rodriguez was one of the ten document custodians it originally selected and he appears on Meta's Rule 26 disclosures as someone Meta represents it may use to support its claims or defenses. Further, until Meta served its amended disclosures on October 21, 2024, Meta's disclosures had instructed Plaintiffs that they could not attempt to contact Mr. Rodriguez except through Meta's outside counsel (Cooley LLP).

Plaintiffs respectfully request that the Court order Meta to provide the last known contact information for these witnesses so Plaintiffs can attempt to obtain documents and deposition testimony from them through the appropriate procedures. Plaintiffs also request that the Court order Meta to add Ms. Zhang, Mr. Lample, and Mr. Grave as document custodians given their importance to the claims and defenses at issue in this case, and because Meta has represented in meeting and conferring with Plaintiffs over the parties' document disputes that Meta has not searched for and produced responsive emails or like communications in other non-custodial systems but only communications in their document custodians' files. That is contrary to this District's ESI Guidelines and general e-discovery practice. *See, e.g.*, N.D. Cal. ESI Discovery Checklist, § III (discovery must be prioritized from systems like email, finance, HR).

Meta's Position:

As a preliminary matter, this request from Plaintiffs is another untimely discovery demand that (1) does not relate to deposition scheduling and (2) relates to a discovery issue for which the District Court entered deadlines for the Parties to meet and confer (October 16) and file discovery motions (October 23). *See* ECF Nos. 227, 238. Plaintiffs' request should be denied for these reasons alone.

Yesterday afternoon, Plaintiffs for the first time told Meta their intention to seek a court order requiring Meta to provide the "last known contact information" for four individuals: Susan Zhang, Guillaume Lample, Edouard Grave, and Aurelian Rodriguez. These are all *former* Meta employees over whom Meta has no control. Plaintiffs cite no authority requiring Meta to provide information relating to individuals outside its control, let alone those located outside of the United States. Separately, Plaintiffs ask the Court to "add" Ms. Zhang, Mr. Lample, and Mr. Grave as document custodians; this request is both untimely and inappropriate.

Plaintiffs mischaracterize the sequence of events leading up to this filing. On September 6, Meta informed Plaintiffs that Mr. Lample, Mr. Grave, and Mr. Rodriguez no longer worked for the company. (Lauter Decl., Ex H). On October 17, after this Court ordered that the parties may take additional depositions (ECF No. 231), Plaintiffs sent Meta a list of witnesses they sought to

depose and asked for their availability, which included Mr. Lample, Mr. Grave, and Mr. Rodriguez. The next day, Meta's counsel again informed Plaintiffs that they do not represent Mr. Lample, Mr. Grave, and Mr. Rodriguez. Then, on October 18, Plaintiffs for the first time requested deposition availability for Susan Zhang, another former Meta employee. Plaintiffs claim that they "repeatedly" asked Meta for the "last known contact information" for Mr. Lample, Mr. Grave, and Mr. Rodriguez. In fact, Plaintiffs first made this request on October 18, before requesting deposition availability for Ms. Zhang, and only raised the issue once more on October 22, asking for Mr. Rodriguez's contact information after Meta served amended Rule 26 disclosures.

Plaintiffs cite no authority for their position that Meta is required to share the last known contact information for former employees no longer within its control, let alone for employees located outside of the United States. Moreover, Plaintiffs have not explained whether they have done anything to try to find contact information themselves. As to Mr. Lample, Mr. Grave, and Mr. Rodriguez, doing so could be in contravention of the European Union's privacy laws, including the General Data Protection Regulation (GDPR), which sets forth detailed requirements for companies for handling personal and employee data.²⁵ Nonetheless, a search on LinkedIn (a platform equally available to Plaintiffs) shows that Mr. Lample is currently employed at Mistral AI,²⁶ Mr. Grave is employed at Kyutai,²⁷ and Mr. Rodriguez is employed at Cohere.²⁸ As for Ms. Zhang, Meta attempted to contact her, but was unsuccessful. There is nothing further for Meta to do in this regard.

Plaintiffs make much of the fact that Mr. Rodriguez was originally listed in Meta's Rule 26 disclosures, which indicated that he should be contacted through Meta, but this does not help them. Mr. Rodriguez left Meta on June 28, 2024, nearly seven months after Meta served its initial disclosures. Although Plaintiffs are correct that Meta served amended initial disclosures on October 21, they leave out that Meta informed Plaintiffs **on September 6** that Mr. Rodriguez no longer worked at the company. Moreover, Meta has no ability to compel this former employee residing in France to appear for deposition, particularly in view of the requirements of French law and the Hague Convention process. Indeed, Plaintiffs cite no authority holding that a party is expected to make a foreign, former employee available simply because they were previously listed on Rule 26 disclosures.

Even if Meta could provide Plaintiffs with contact information for Mr. Lample, Mr. Grave, and Mr. Rodriguez, doing so would be futile. Because these individuals reside in France, in order to seek discovery from them, Plaintiffs must follow the procedures outlined under the Hague Convention. This is not news to Plaintiff; many weeks ago, in previously seeking to enlarge the discovery period, Plaintiffs specifically cited the need to follow the lengthy procedures under the Hague Convention to depose former employees located outside of the United States. *See* ECF No. 129 at 1, 3, 4. With the close of discovery just over six weeks away, deposing Mr. Lample, Mr.

²⁵ Given the compressed timing and Plaintiffs' failure to provide any authority for their request, this issue would need to be the subject of more thorough research and briefing to ensure that all parties are complying with European Union law and/or other applicable US law.

²⁶ <https://www.linkedin.com/in/guillaume-lample-7821095b/?originalSubdomain=fr>

²⁷ <https://www.linkedin.com/in/edouard-grave-63099823/?originalSubdomain=fr>

²⁸ <https://www.linkedin.com/in/aur%C3%A9lien-rodriguez-145b75134/?originalSubdomain=fr>

Grave, or Mr. Rodriguez within the discovery period will be all but impossible, particularly because the US Embassy requires at least 45 days notice to request permission for a voluntary deposition from the French authorities.²⁹ Moreover, to the extent Plaintiffs intend to “attempt to obtain documents” from these individuals, they are well past the October 18 deadline to serve written discovery. ECF No. 238.³⁰

Further, Plaintiffs’ request is also improper because they are out of depositions. In its October 15, 2024 Order (ECF No. 231), the Court allowed Plaintiffs to take 21 additional depositions (beyond the four it had taken), excluding Meta’s Rule 30(b)(6) witness. Since then, Plaintiffs have sought deposition availability for at least **24** additional witnesses, and the parties have scheduled 19 depositions, which include the 30(b)(6) designees. *See infra* Section I & II. Indeed, when Meta informed Plaintiffs that it did not represent Mr. Lample, Mr. Grave and Mr. Rodriguez, Plaintiffs swapped their names for other deponents. (Lauter Decl., Ex. I at 5.) Excluding Plaintiffs’ demand for Mr. Cox, Plaintiff has only 1 more deposition left to be scheduled. Plaintiffs therefore **cannot** depose each of Ms. Zhang, Mr. Lample, Mr. Grave, and Mr. Rodriguez without exceeding the Court’s prior limit of 21 additional depositions.

Finally, it is inappropriate for Plaintiffs to ask the Court – in this briefing, and in contravention of the parties’ Stipulated Order Re: Discovery of Electronically Stored Information (ECF No. 101) – to “order Meta to add Ms. Zhang, Mr. Lample, and Mr. Grave as document custodians.” First, there is an ESI protocol for a reason, and raising this issue at the end of a last-minute drop-in is improper. Second, as noted above, this joint letter brief is intended to address outstanding issues concerning depositions. ECF No. 243. Adding document custodians is outside of that scope. Third, the Parties are now less than six weeks out from the close of fact discovery, and adding ever more ESI custodians will be yet another “major change[] to the scope of discovery.” ECF 196 at 2 (“But eleventh hour motions to compel should be for wrap-up items, not for major changes to the scope of discovery. Adding five document custodians is a big change, not a small change, to the scope of Meta’s document production obligations in this case.”) Moreover, Plaintiffs have waived any demands for additional ESI custodians, as those demands were not the subject of Letter Briefs on Existing Written Discovery that were due on October 23.³¹ Plaintiffs’ request should be denied.

²⁹ France Judicial Assistance Information, <https://travel.state.gov/content/travel/en/legal/Judicial-Assistance-Country-Information/France.html> (“The U.S. Embassy or Consulate must have the documentation at least 45 days prior to the proposed deposition date and French translations in order to request permission of the French Central Authority. ”)

³⁰ Plaintiffs have disregarded this deadline on multiple occasions, serving two sets of requests for production on October 23 and 26, respectively, fourteen Rule 45 subpoenas on October 25, and two more 45 subpoenas on October 28 and 30, respectively. Meta has written to Plaintiffs regarding this untimely written discovery and reserves all rights in that regard.

³¹ Plaintiffs refer to representations allegedly made by Meta regarding Meta’s document collection efforts, which Meta does not understand (“Meta has not searched for and produced responsive emails or like communications in other non-custodial systems but only communications in their document custodians’ files.”). In any event, this letter brief is intended to discuss issues relating to depositions, not document collection efforts, which were not timely raised by Plaintiffs in October 23 Letter Briefs on Existing Written Discovery.

/s/Maxwell Pritt

Maxwell Pritt

Reed Forbush

Boies Schiller Flexner, LLP

Attorneys for Plaintiffs

/s/Bobby Ghajar

Bobby Ghajar

Cooley LLP

Attorneys for Defendant

ATTESTATION PURSUANT TO CIVIL LOCAL RULE 5-1(h)

I hereby attest that I obtained concurrence in the filing of this document from each of the other signatories. I declare under penalty of perjury that the foregoing is true and correct.

Dated: October 31, 2024

/s/Bobby Ghajar

Bobby Ghajar

Cooley LLP

Attorneys for Defendant